M. Krasnov A. Kiselev
G. Makarenko E. Shikin

Mathematical Analysis for Engineers

Volume 2

Mir Publishers Moscow

Курс
высшей математики
для инженеров

М. Краснов,
А. Киселев,
Г. Макаренко,
Е. Шикин

В двух томах

Том
2

Mathematical Analysis for Engineers

M. Krasnov
A. Kiselev
G. Makarenko
E. Shikin

In two volumes

Volume
2

Mir Publishers
Moscow

Translated from Russian
by Alexander Repyev

First published 1990

На английском языке

Printed in the Union of Soviet Socialist Republics

ISBN 5-03-000271-5
ISBN 5-03-000269-3

Contents

Preface

This two-volume book was written for students of technical colleges who have had the usual mathematical training. It contains just enough information to continue with a wide variety of engineering disciplines. It covers analytic geometry and linear algebra, differential and integral calculus for functions of one and more variables, vector analysis, numerical and functional series (including Fourier series), ordinary differential equations, functions of a complex variable, Laplace and Fourier transforms, and equations of mathematical physics. This list itself demonstrates that the book covers the material for both a basic course in higher mathematics and several special sections that are important for applied problems. Hence, it may be used by a wide range of readers. Besides students in technical colleges and those starting a mathematics course, it may be found useful by engineers and scientists who wish to refresh their knowledge of some aspects of mathematics.

We tried to give the fundamental material concisely and without distracting detail. We concentrated on the prensentation of the basic ideas of linear algebra and analysis to make it detailed and as comprehensible as possible. Mastery of these ideas is a requirement to understand the later material.

The many examples also serve this aim. The examples were written to help students with the mechanics of solving typical problems.

More than 600 diagrams are simple illustrations, clear enough to demonstrate the ideas and statements convincingly, and can be fairly easily reproduced.

We were conscious not to burden the course with scrupulous proofs for theorems which have little practical application. As a rule we chose the proof (marked in the text with special symbols) that was constructive in nature or explained fundamental ideas that had been introduced, showing how they work. This approach made it possible to devise algorithms for solving whole classes of important problems.

In addition to the examples, we have included several carefully selected problems and exercises (around 1000) which should be of interest to those pursuing an independent mathematics course. The problems have the form

of moderately sized theorems. They are very simple but are good training for those learning the fundamental ideas.

Chapters 1-6, 26 and Appendix II were written by E.Shikin, Chapters 7-8, 11, 12, 17-21, 27, 28 and 29-32 by M.Krasnov, Chapters 9, 10, 13-16 by A.Kiselev, and Chapters 22-25 and Appendix I by G.Makarenko. There was no general editor, but each of the authors has read the chapters written by the colleagues, and so each chapter benefited from collective advice.

The Authors

Chapter 13

Number Series

13.1 Definition. Sum of a Series

Consider an infinite number sequence

$a_1,\ a_2,\ \ldots,\ a_n,\ \ldots\ .$

A number series is an expression of the form

$$a_1 + a_2 + \ldots + a_n + \ldots\ . \tag{13.1}$$

A shorthand notation for this is $\sum_{n=1}^{\infty} a_n$.

The numbers $a_1, a_2, \ldots$ are called the *terms* of the series, and the number a_n is called the *nth* or *general term* of the series.

The sum of the finite number n of the terms of the series is called the *nth partial sum* of the series:

$$S_n = a_1 + a_2 + \ldots + a_n = \sum_{k=1}^{n} a_k\ .$$

Now consider the sequence $\{S_n\}$ of partial sums of the series (13.1)

$$S_1 = a_1,$$

$$S_2 = a_1 + a_2,$$

$$\vdots$$

$$S_n = a_1 + a_2 + \ldots + a_n,\ \ldots\ .$$

Definition. If the sequence $\{S_n\}$ has a finite limit $\lim_{n\to\infty} S_n = S$, i.e., $\{S_n\}$ converges to S, then the limit is called the *sum* of the series $\sum_{n=1}^{\infty} a_n$ and the series is said to *converge*: $\sum_{n=1}^{\infty} a_n = S$. If there is no $\lim_{n\to\infty} S_n$, i.e., $\{S_n\}$ does not converge, then the series $\sum_{n=1}^{\infty} a_n$ is said to *diverge* and to have no sum.

Examples. (1) Show that the series

$$\frac{1}{3} + \frac{1}{15} + \frac{1}{35} + \dots + \frac{1}{4n^2 - 1} + \dots = \sum_{n=1}^{\infty} \frac{1}{4n^2 - 1}$$

converges.

◀ We consider the nth partial sum of the series

$$S_n = \frac{1}{3} + \frac{1}{15} + \frac{1}{35} + \dots + \frac{1}{4n^2 - 1}.$$

Using the obvious relation

$$\frac{1}{4n^2 - 1} = \frac{1}{(2n - 1)(2n + 1)} = \frac{1}{2}\left(\frac{1}{2n - 1} - \frac{1}{2n + 1}\right)$$

we represent S_n in the form

$$S_n = \frac{1}{1 \times 3} + \frac{1}{3 \times 5} + \frac{1}{5 \times 7} + \dots + \frac{1}{(2n - 1)(2n + 1)}$$
$$= \frac{1}{2}\left(1 - \frac{1}{3}\right) + \frac{1}{2}\left(\frac{1}{3} - \frac{1}{5}\right) + \frac{1}{2}\left(\frac{1}{5} - \frac{1}{7}\right)$$
$$+ \dots + \frac{1}{2}\left(\frac{1}{2n - 1} - \frac{1}{2n + 1}\right) = \frac{1}{2}\left(1 - \frac{1}{2n + 1}\right).$$

Passing to the limit as $n \to \infty$, we will have

$$\lim_{n \to \infty} S_n = \frac{1}{2}.$$

By the definition the series converges and its sum is $S = 1/2$, or

$$\sum_{n=1}^{\infty} \frac{1}{4n^2 - 1} = \frac{1}{2}. \blacktriangleright$$

(2) Consider a series known as a geometric progression with a ratio q

$$a + aq + aq^2 + \dots + aq^{n-1} + \dots = \sum_{n=1}^{\infty} aq^{n-1} \qquad (a \neq 0).$$

◀ The sum of n terms is,

$$S_n = a + aq + aq^2 + \dots + aq^{n-1}$$
$$= \frac{a - aq^n}{1 - q} = \frac{a}{1 - q} - \frac{aq^n}{1 - q} \qquad (q \neq 1).$$

If $|q| < 1$, then $\lim_{n\to\infty} q^n = 0$, and so

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty}\left(\frac{a}{1-q} - \frac{aq^n}{1-q}\right) = \frac{a}{1-q},$$

i.e., the series converges and its sum is $\frac{a}{1-q}$, or

$$\sum_{n=1}^{\infty} aq^{n-1} = \frac{a}{1-q}.$$

If $|q| > 1$, then $\lim_{n\to\infty} q^n = \infty$ and hence $\lim_{n\to\infty} S_n = \infty$, i.e., the series diverges.

At $q = -1$ we obtain the divergent series $a - a + a - a + \ldots$ $(a \neq 0)$. Its partial sum is

$$S_n = \begin{cases} a & \text{for odd } n, \\ 0 & \text{for even } n. \end{cases}$$

It follows that $\lim_{n\to\infty} S_n$ does not exist.

At $q = 1$ we will have the series $a + a + a + \ldots$, for which $S_n = na$, and hence $\lim_{n\to\infty} S_n = \lim_{n\to\infty} na = \infty$, i.e., the series diverges.

Consequently, the series $a + aq + aq^2 + \ldots + aq^{n-1} + \ldots$ converges for $|q| < 1$, its sum being $\frac{a}{1-q}$, and diverges for $|q| \geqslant 1$. ►

13.2 Operations on Series

Operations on number series may be deduced from the following theorems:

Theorem 13.1. *If the series* $a_1 + a_2 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n$ *converges, then so does the series obtained from it by discarding any finite number of terms in the beginning. Conversely, if a series obtained from the given series by discarding a finite number of terms in the beginning converges, then the given series converges.*

◄ In the partial sum

$$S_n = a_1 + a_2 + \ldots + a_k + a_{k+1} + \ldots + a_n$$

of the series we denote by σ_k the sum of the first $k (k < n)$ discarded terms. We get

$$S_n = (a_1 + a_2 + \ldots + a_k) + (a_{k+1} + \ldots + a_n) = \sigma_k + S_{n-k}.$$

It follows that if there exists $\lim\limits_{n\to\infty} S_n$, then there exists $\lim\limits_{n\to\infty} S_{n-k}$ (k = const) and, conversely, if there exists $\lim\limits_{n\to\infty} S_{n-k}$, then there exists $\lim\limits_{n\to\infty} S_n$. ►

Remark. The resultant series $a_{k+1} + a_{k+2} + \dots$ has the sum $\tilde{S} = S - \sigma_k$, S being the sum of the original series.

Theorem 13.2. *Let the series* $a_1 + a_2 + \dots + a_n + \dots = \sum\limits_{n=1}^{\infty} a_n$ *be convergent and* $\lambda \neq 0$ *be a number. Then the series*

$$\lambda a_1 + \lambda a_2 + \dots + \lambda a_n + \dots = \sum_{n=1}^{\infty} \lambda a_n$$

converges and

$$\sum_{n=1}^{\infty} \lambda a_n = \lambda \sum_{n=1}^{\infty} a_n.$$

◄ We write partial sums for the series $\sum\limits_{n=1}^{\infty} a_n$ and $\sum\limits_{n=1}^{\infty} \lambda a_n$:

$$S_n = a_1 + a_2 + \dots + a_n, \qquad \sigma_n = \lambda a_1 + \lambda a_2 + \dots + \lambda a_n.$$

Clearly, $\sigma_n = \lambda S_n$. Since, as stated, the series $\sum\limits_{n=1}^{\infty} a_n$ converges, i.e., there exists $\lim\limits_{n\to\infty} S_n$, then from the last equality there is $\lim\limits_{n\to\infty} \sigma_n$, such that $\lim\limits_{n\to\infty} \sigma_n = \lim\limits_{n\to\infty} \lambda S_n = \lambda \lim\limits_{n\to\infty} S_n$, i.e.,

$$\sum_{n=1}^{\infty} \lambda a_n = \lambda \sum_{n=1}^{\infty} a_n. \text{ ►}$$

Theorem 13.3. *If the series* $\sum\limits_{n=1}^{\infty} a_n$ *and* $\sum\limits_{n=1}^{\infty} b_n$ *converge, then their sum and difference, i.e.,* $\sum\limits_{n=1}^{\infty} (a_n + b_n)$ *and* $\sum\limits_{n=1}^{\infty} (a_n - b_n)$, *converge, and*

$$\sum_{n=1}^{\infty} (a_n \pm b_n) = \sum_{n=1}^{\infty} a_n \pm \sum_{n=1}^{\infty} b_n.$$

◄ Let $S_n = a_1 + a_2 + \dots + a_n$, $\tilde{S}_n = b_1 + b_2 + \dots + b_n$ and $\sigma_n = (a_1 + b_1) + (a_2 + b_2) + \dots + (a_n + b_n)$ be partial sums of the series

$$\sum_{n=1}^{\infty} a_n, \quad \sum_{n=1}^{\infty} b_n, \quad \sum_{n=1}^{\infty} (a_n + b_n),$$

respectively. Clearly, $\sigma_n = S_n + \tilde{S}_n$. Since by the statement of the theorem the series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converge, i.e., there exist $\lim_{n\to\infty} S_n$ and $\lim_{n\to\infty} \tilde{S}_n$, it follows from the last equality, which holds for all n, that there exists $\lim_{n\to\infty} \sigma_n$, and that

$$\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} (S_n + \tilde{S}_n) = \lim_{n\to\infty} S_n + \lim_{n\to\infty} \tilde{S}_n,$$

which is equivalent to

$$\sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n. \blacktriangleright$$

In a like manner, we can prove the convergence of $\sum_{n=1}^{\infty} (a_n - b_n)$.

We now introduce the concept of the remainder of a series, which we will use later.

Definition. If we discard the first n terms in the convergent series

$$a_1 + a_2 + \dots + a_n + a_{n+1} + a_{n+2} + \dots = \sum_{n=1}^{\infty} a_n$$

we will obtain the convergent series

$$a_{n+1} + a_{n+2} + \dots + a_{n+k} + \dots = \sum_{k=1}^{\infty} a_{n+k},$$

which is called the nth *remainder* of the series and denoted by

$$R_n = \sum_{k=1}^{\infty} a_{n+k}$$

for a fixed n.

The original series can then be written as

$$\sum_{n=1}^{\infty} a_n = S_n + R_n.$$

If S is the sum of the series $\sum_{n=1}^{\infty} a_n$, then the remainder will be $R_n = S - S_n$ for any $n = 1, 2, \dots$.

For example, for the series $\sum_{n=1}^{\infty} aq^{n-1}$ its nth remainder will be the series

$$R_n = aq^n + aq^{n-1} + \dots + aq^{n+k-1} + \dots = \sum_{k=1}^{\infty} aq^{k+n-1},$$

which is convergent for $|q| < 1$.

13.3 Tests for Convergence of Series

Cauchy criterion for convergence of a sequence allows to formulate the general criterion for convergence of a number series.

Theorem 13.4 (Cauchy criterion). *A necessary and sufficient condition for a number series* $\sum_{n=1}^{\infty} a_n$ *to converge is that for any number* $\varepsilon > 0$ *there exists a number* $N = N(\varepsilon)$ *such that for any* $n > N$ *there holds*

$$|a_n + a_{n+1} + \dots + a_{n+p}| < \varepsilon \tag{13.2}$$

for all $p = 0, 1, 2, \dots$.

In terms of the partial sums S_{n+p} and S_{n-1} of $\sum_{n=1}^{\infty} a_n$, we can write (13.2) as

$$|S_{n+p} - S_{n-1}| < \varepsilon.$$

From the Cauchy criterion follows the *necessary test for convergence of a number series.*

Theorem 13.5. *If series* $\sum_{n=1}^{\infty} a_n$ *converges, then* $\lim_{n\to\infty} a_n = 0$.

◄ Putting $p = 0$ in Theorem 13.4, we will have $|a_n| < \varepsilon$, which holds for all $n > N(\varepsilon)$. The number $\varepsilon > 0$ being arbitrary, we have $\lim_{n\to\infty} a_n = 0$. ►

Corollary. If $\lim_{n\to\infty} a_n$ is not equal to zero or does not exist, then $\sum_{n=1}^{\infty} a_n$ diverges.

◄ Suppose that $\sum_{n=1}^{\infty} a_n$ converges. Then by Theorem 13.4, there must exist $\lim_{n\to\infty} a_n$ equal to zero. Our assumption has led us to a contradiction, hence it is wrong. Therefore, the series is divergent. ►

Examples. (1) The number series

$$-1 + 0 + \frac{1}{2} + \frac{\sqrt{2}}{2} + \cos\frac{\pi}{5} + \dots = \sum_{n=1}^{\infty} \cos\frac{\pi}{n}$$

diverges, since

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \cos\frac{\pi}{n} = \cos 0 = 1 \neq 0.$$

(2) The series

$$1 - 1 + 1 - 1 + \ldots = \sum_{n=1}^{\infty} (-1)^{n+1}$$

diverges, since $\lim_{n\to\infty} a_n = \lim_{n\to\infty} (-1)^{n+1}$ does not exist.

Remark. Theorem 13.5 gives a necessary condition for a series to converge, which is not, however, sufficient, i.e., the condition $\lim_{n\to\infty} a_n = 0$ may be met for a divergent series $\sum_{n=1}^{\infty} a_n$ as well.

(3) Consider the number series

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n},$$

known as the *harmonic series*. The harmonic series meets the necessary condition, since

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{1}{n} = 0.$$

We will prove that the series is divergent, for which purpose we will make use of the Cauchy criterion. Putting $p = n$ gives

$$a_n + a_{n+1} + a_{n+2} + \ldots + a_{2n} = \frac{1}{n} + \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n}$$

$$> \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} > \frac{1}{n+n} + \frac{1}{n+n} + \ldots + \frac{1}{2n}$$

$$= \frac{1}{2n} + \frac{1}{2n} + \ldots + \frac{1}{2n} = n \cdot \frac{1}{2n} = \frac{1}{2}.$$

The inequality holds for any arbitrarily large n. It follows that for $\varepsilon \leqslant 1/2$ and $p = n$ inequality (13.2) is not valid, and by the Cauchy criterion the harmonic series is divergent.

Comparison tests for series of positive terms make it possible to establish whether or not a number series is convergent (divergent) by comparing it with another series that is known to be convergent (divergent).

Theorem 13.6. *Let*

$$a_1 + a_2 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n, \tag{13.3}$$

$$b_1 + b_2 + \ldots + b_n + \ldots = \sum_{n=1}^{\infty} b_n \tag{13.4}$$

be two series of positive terms, such that

$$a_n \leqslant b_n \tag{13.5}$$

for all n. Then, if $\sum_{n=1}^{\infty} b_n$ *converges,* $\sum_{n=1}^{\infty} a_n$ *converges as well; and if* $\sum_{n=1}^{\infty} a_n$ *diverges, then* $\sum_{n=1}^{\infty} b_n$ *diverges as well.*

◄ We form the partial sums of (13.3) and (13.4)

$$S_n = a_1 + a_2 + \dots + a_n; \qquad \sigma_n = b_1 + b_2 + \dots + b_n.$$

It follows from (13.5) that $S_n \leqslant \sigma_n$ for all $n = 1, 2, \dots$.

(1) Suppose that series (13.4) converges, i.e., there exists $\lim_{n\to\infty} \sigma_n = \sigma$ of its partial sum. Since the terms of these series are positive, then $0 < \sigma_n < \sigma$, and it follows by (13.5) that $0 < S_n < \sigma$ for $n = 1, 2, \dots$. Thus, all the partial sums S_n of (13.3) are bounded and increase with n, since $a_n > 0$ for all n. Consequently, the sequence of partial sums $\{S_n\}$ is convergent, i.e., there exists $\lim_{n\to\infty} S_n = S$, which implies that $\sum_{n=1}^{\infty} a_n$ is a convergent series. Now, from the inequality $0 < S_n < \sigma$, which holds for all natural n, we obtain as $n \to \infty$ the inequality $0 \leqslant S \leqslant \sigma$, i.e., the sum S of (13.3) does not exceed the sum σ of the convergent series (13.4).

(2) Suppose that $\sum_{n=1}^{\infty} a_n$ diverges. Since all $a_n > 0$, then S_n increases with n, and hence $\lim_{n\to\infty} S_n = +\infty$. From the inequality $\sigma_n \geqslant S_n$ $(n = 1, 2, \dots)$ we get $\lim_{n\to\infty} \sigma_n = +\infty$, i.e., $\sum_{n=1}^{\infty} b_n$ diverges.

Remark 1. The theorem is valid even when (13.5) holds not for all n, but only beginning with some k, i.e., for all $n \geqslant k$, since when we drop a finite number of terms in the beginning, we do not violate the convergence of the series.

Examples. (1) Examine for convergence the series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2^n + \sqrt{n}}$.

◄ We have

$$\frac{1}{2^n + \sqrt{n}} \leqslant \frac{1}{2^n} = \left(\frac{1}{2}\right)^n \qquad (n = 0, 1, 2, \dots).$$

Since $\sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n$ converges, then by the comparison test the original series converges as well.

(2) Examine for convergence the series $\sum_{n=2}^{\infty}\frac{1}{\ln n}$.

◄ From the inequality $\ln n < n$ follows $1/\ln n > 1/n$ for $n = 2, 3, \ldots$, and the harmonic series $\sum_{n=1}^{\infty}\frac{1}{n}$ diverges $\left(\text{in this case } \sum_{n=2}^{\infty}\frac{1}{n} \text{ diverges too}\right)$. Then by the comparison test the original series diverges. ►

Remark 2. Theorem 13.6 is also valid for a more general inequality $a_n \leqslant \lambda b_n$ $(\lambda a_n \leqslant b_n)$ $(n = 1, 2, \ldots)$, where $\lambda > 0$.

(3) Examine for convergence the series $\sum_{n=1}^{\infty}\left(1 - \cos\frac{\pi}{2^n}\right)$.

◄ Using the inequality $\sin x \leqslant x$, which holds for all $x \geqslant 0$, we find that

$$0 < 1 - \cos\frac{\pi}{2^n} = 2\sin^2\frac{\pi}{2 \times 2^n} \leqslant 2\left(\frac{\pi}{2 \times 2^n}\right)^2 = \frac{\pi^2}{2}\frac{1}{4^n}$$

for $n = 1, 2, \ldots$. Here $\lambda = \pi^2/2$, and $\sum_{n=1}^{\infty}\frac{1}{4^n}$ converges (see Sec. 13.1, Example 2). By comparison test (and considering Remark 2), the original series converges. ►

From Theorem 13.6 follows a corollary.

Corollary. If there exists a finite nonzero

$$\lim_{n\to\infty}\frac{a_n}{b_n} = L \quad (0 < L < +\infty),$$

then (13.3) and (13.4) converge or diverge simultaneously.

◄ Indeed, from the existence of the above limit it follows that for any $\varepsilon > 0$ subject to the condition $L - \varepsilon > 0$, there exists a number N such that for all $n > N$

$$\left|\frac{a_n}{b_n} - L\right| < \varepsilon \quad \text{or} \quad L - \varepsilon < \frac{a_n}{b_n} < L + \varepsilon.$$

Hence

$$(L - \varepsilon)\, b_n < a_n < (L + \varepsilon)\, b_n \quad \text{for} \quad n > N.$$

If (13.4) converges, so does $\sum_{n=1}^{\infty}(L+\varepsilon)b_n$. But since $a_n < (L+\varepsilon)b_n$ for $n > N$, then, by Theorem 13.6 (and Remark 2), series (13.3) will converge as well. If (13.4) diverges, then $\sum_{n=1}^{\infty}(L-\varepsilon)b_n$ diverges as well. And since $(L-\varepsilon)b_n < a_n$ for $n > N$, then, by Theorem 13.6, series (13.3) also diverges. ►

Examples. (1) Examine for convergence the series $\sum_{n=1}^{\infty}\sin\frac{\pi}{n}$.

◄ Compare the given series with the harmonic series $\sum_{n=1}^{\infty}\frac{1}{n}$. We have

$$\lim_{n\to\infty}\frac{a_n}{b_n} = \lim_{n\to\infty}\frac{\sin\frac{\pi}{n}}{\frac{1}{n}} = \lim_{n\to\infty}\pi\frac{\sin\frac{\pi}{n}}{\frac{\pi}{n}} = \pi \neq 0.$$

The harmonic series diverges, therefore, by the corollary, the original series diverges too. ►

(2) Examine for convergence the series $\sum_{n=1}^{\infty}\frac{1}{2^n - n}$.

◄ For comparison we take the convergent series $\sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n$. Then

$$\lim_{n\to\infty}\frac{a_n}{b_n} = \lim_{n\to\infty}\frac{\frac{1}{2^n-n}}{\frac{1}{2^n}} = \lim_{n\to\infty}\frac{2^n}{2^n-n} = \lim_{n\to\infty}\frac{1}{1-\frac{n}{2^n}} = 1 \neq 0,$$

since $\lim_{n\to\infty}\frac{n}{2^n} = 0$. Therefore, the original series converges. ►

D'Alembert test for convergence of a series is given by the following theorem.

Theorem 13.7. *Consider* $\sum_{n=1}^{\infty}a_n$, *where* $a_n > 0$. *If there exists*

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = \lambda,$$

then for $0 \leqslant \lambda < 1$ *the series converges, and for* $\lambda > 1$ *the series diverges.*

◄ Suppose that there exists the finite limit

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lambda,$$

where $\lambda < 1$. Take any number q such that $\lambda < q < 1$. Then, for any $\varepsilon > 0$, e.g., for $\varepsilon = q - \lambda$, there exists a number N such that for all $n \geqslant N$ we will have

$$\left| \frac{a_{n+1}}{a_n} - \lambda \right| < q - \lambda.$$

Specifically, we will have

$$\frac{a_{n+1}}{a_n} - \lambda < q - \lambda \quad \text{or} \quad \frac{a_{n+1}}{a_n} < q.$$

Hence $a_{n+1} < a_n q$ for all $n \geqslant N$. From the latter inequality we will obtain, letting n assume the values $N, N+1, N+2, \ldots$

$$a_{N+1} < a_N q,$$
$$a_{N+2} < a_{N+1} q < a_N q^2,$$
$$a_{N+3} < a_{N+2} q < a_N q^3,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .$$

So, the terms of the series

$$a_{N+1} + a_{N+2} + a_{N+3} + \ldots$$

are not larger than the corresponding terms of the series

$$a_N q + a_N q^2 + a_N q^3 + \ldots,$$

which converges as a series of the terms of a geometric progression with the ratio q, where $0 < q < 1$. By comparison test the series $a_{N+1} + a_{N+2} + a_{N+3} + \ldots$ converges, and hence $\sum_{n=1}^{\infty} a_n$ converges too.

If $\lambda > 1$, then beginning with a certain number N we will have

$$\frac{a_{N+1}}{a_N} > 1 \quad \text{or} \quad a_{N+1} > a_N > 0.$$

Consequently, $\lim_{n\to\infty} a_n \neq 0$ and the series $\sum_{n=1}^{\infty} a_n$ diverges, since the required test for convergence is not satisfied.

Remark. If $\lim\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n} = 1$, then the D'Alembert test gives no answer as to whether or not a series converges.

Examples. (1) Examine for convergence the series $\sum\limits_{n=1}^{\infty} \dfrac{n^2}{2^n}$.

◄ We have

$$a_n = \frac{n^2}{2^n} \quad \text{and} \quad a_{n+1} = \frac{(n+1)^2}{2^{n+1}}.$$

Then

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{(n+1)^2 2^n}{2^{n+1} n^2} = \lim_{n\to\infty} \frac{1}{2}\left(1 + \frac{1}{n}\right)^2 = \frac{1}{2} < 1.$$

By the D'Alembert test, the series converges. ►

(2) Examine for convergence the series $\sum\limits_{n=1}^{\infty} \dfrac{n^n}{n!}$.

◄ We have

$$a_n = \frac{n^n}{n!}, \quad a_{n+1} = \frac{(n+1)^{n+1}}{(n+1)!} = \frac{(n+1)^n}{n!};$$

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{(n+1)^n n!}{n!\, n^n} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e > 1.$$

The series diverges. ►

Cauchy test. Theorem 13.8. *Consider the series* $\sum\limits_{n=1}^{\infty} a_n$, $a_n > 0$, $n = 1, 2, \ldots$. *If there exists a finite* $\lim\limits_{n\to\infty} \sqrt[n]{a_n} = \lambda$, *then for* $0 \leqslant \lambda < 1$ *the series converges, and for* $\lambda > 1$ *diverges.*

◄ (1) Let $\lambda < 1$. We take q such that $\lambda < q < 1$. Since there exists $\lim\limits_{n\to\infty} \sqrt[n]{a_n} = \lambda$, where $\lambda < q$, then beginning with a certain number N we will have $\sqrt[n]{a_n} < q$, whence $a_n < q^n$ for $n \geqslant N$. And so the terms from a_{N+1} on are smaller than the corresponding terms of the convergent series $\sum\limits_{n=N}^{\infty} q^n$. By the comparison test $\sum\limits_{n=N}^{\infty} a_n$ converges, and hence $\sum\limits_{n=1}^{\infty} a_n$ converges too.

(2) Let $\lambda > 1$. Then, beginning with a certain N and for all $n > N$, we will have $\sqrt[n]{a_n} > 1$ or $a_n > 1$. Accordingly, $\lim\limits_{n\to\infty} a_n \neq 0$ and the series $\sum\limits_{n=1}^{\infty} a_n$ diverges.

Remark. If $\lambda = 1$, then $\sum_{n=1}^{\infty} a_n$ can either converge or diverge.

Examples. (1) Examine for convergence the series $\sum_{n=1}^{\infty} \frac{2^n}{\ln^n (n+1)}$.

◄ We have

$$a_n = \frac{2^n}{\ln^n (n+1)}, \quad \sqrt[n]{a_n} = \frac{2}{\ln (n+1)},$$

$$\lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \frac{2}{\ln (1+n)} = 0 < 1.$$

The series converges. ►

(2) Examine for convergence the series $\sum_{n=1}^{\infty} \frac{1}{2^n}\left(1 + \frac{1}{n}\right)^{n^2}$.

◄ Here

$$a_n = \frac{1}{2^n}\left(1 + \frac{1}{n}\right)^{n^2}, \quad \sqrt[n]{a_n} = \frac{1}{2}\left(1 + \frac{1}{n}\right)^{n},$$

$$\lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \frac{1}{2}\left(1 + \frac{1}{n}\right)^{n} = \frac{1}{2}\,\mathrm{e} > 1.$$

The series diverges. ►

Integral test. Theorem 13.9. *Let a function $f(x)$ be defined, continuous, positive and nonincreasing for $x \geqslant 1$. Then the number series $\sum_{n=1}^{\infty} f(n)$ converges, if the integral $\int_1^{+\infty} f(x)\,dx$ converges; and diverges, if the integral diverges.*

◄ We plot $f(x)$ and take points with $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, ..., $x_n = n$ on the curve.

We now consider two stepped figures shown in Fig. 13.1.

The area Q of the curvilinear trapezoid bounded by the straight lines $x = 1$, $x = n$, $y = 0$, and the curve $y = f(x)$ will be

$$Q = \int_1^n f(x)\,dx.$$

We then take the nth partial sum of the series

$$S_n = f(1) + f(2) + f(3) + \dots + f(n).$$

The area of the larger figure will be

$$\bar{Q} = f(2) + f(3) + \ldots + f(n) = S_n - f(1)$$

and the area of the smaller figure will be

$$\bar{\bar{Q}} = f(1) + f(2) + f(3) + \ldots + f(n-1) = S_{n-1}.$$

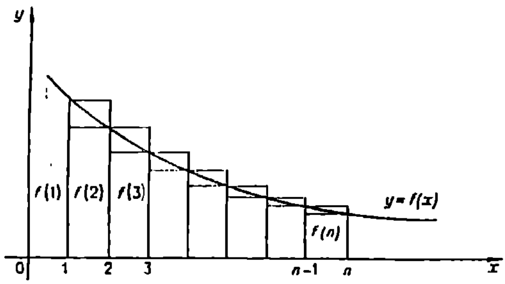


Fig. 13.1

It is seen that $\bar{Q} < Q < \bar{\bar{Q}}$, i.e.,

$$S_n - f(1) < \int_1^n f(x)\,dx < S_{n-1}$$

or

$$S_n - f(1) < \int_1^n f(x)\,dx < S_n \qquad \text{for} \quad n = 1, 2, \ldots \tag{13.6}$$

since $S_{n-1} < S_n$ by virtue of the condition $f(n) > 0$.

(1) Suppose that $\int_1^{+\infty} f(x)\,dx$ converges. Then there will exist $\lim\limits_{n\to\infty} \int_1^n f(x)dx = A$ such that

$$\int_1^n f(x)\,dx \leqslant \int_1^{+\infty} f(x)\,dx = A$$

by virtue of the condition $f(x) > 0$ for $x \in [1, +\infty)$.

It follows from (13.6) that

$$S_n < f(1) + \int_1^n f(x)\,dx \leqslant f(1) + A = M = \text{const},$$

i.e., $0 < S_n < M$ for $n = 1, 2, \ldots$. This means that the sequence $\{S_n\}$ is bounded. The sum S_n increases with n, since $f(n) > 0$ for $n = 1, 2, \ldots$.

Thus, the sequence $\{S_n\}$ of partial sums of the series is strictly monotonous and bounded, and therefore it has $\lim_{n\to\infty} S_n = S$ (see Theorem 7.9), which means that $\sum_{n=1}^{\infty} f(n)$ is convergent.

(2) Suppose now that $\int_1^{+\infty} f(x)\,dx$ diverges. Since $f(x) > 0$ for $x \geqslant 1$, then

$$\int_1^{+\infty} f(x)\,dx = \lim_{n\to\infty} \int_1^n f(x)\,dx = +\infty.$$

It follows from

$$S_n \geqslant \int_1^n f(x)\,dx \quad (n = 1, 2, \ldots)$$

that $\lim_{n\to\infty} S_n = +\infty$, i.e., $\sum_{n=1}^{\infty} f(n)$ diverges. ►

Remark. The theorem also holds for $x \geqslant a$, where a is any number larger than unity.

Examples. (1) Examine for convergence $\sum_{n=1}^{\infty} \frac{1}{n^p}$.

◄ Here $f(n) = 1/n^p$. The integral $\int_1^{+\infty} \frac{dx}{x^p}$ is known (Chap. 11) to converge for $p > 1$ and diverge for $p \leqslant 1$. Hence, the series converges when $p > 1$ and diverges when $p \leqslant 1$. Specifically, at $p = 1$ we will have the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots = \sum_{n=1}^{\infty} \frac{1}{n}$, which has already been shown to diverge. ►

(2) Examine for convergence the series $\sum_{n=1}^{\infty} \frac{1}{n^2 + 1}$.

◄ In this case $f(x) = 1/(x^2 + 1)$. The integral

$$\int_1^{+\infty} \frac{dx}{x^2 + 1} = \lim_{b\to+\infty} \int_1^b \frac{dx}{x^2 + 1} = \lim_{b\to+\infty} \tan^{-1} x \Big|_1^b$$

$$= \lim_{b\to+\infty} (\tan^{-1} b - \tan^{-1} 1) = \frac{\pi}{2} - \frac{\pi}{4} = \frac{\pi}{4},$$

converges, and so does the series. ►

(3) Examine for convergence the series $\sum_{n=1}^{\infty} \frac{n}{n^2+1}$.

◄ Since the nth term of the series is $f(n) = \frac{n}{n^2+1}$ we choose the function $f(x) = \frac{x}{x^2+1}$ satisfying the conditions of the theorem, and then consider the improper integral $\int_1^{+\infty} \frac{x\,dx}{x^2+1}$.

We have

$$\int_1^{+\infty} \frac{x\,dx}{x^2+1} = \lim_{b\to+\infty} \int_1^b \frac{x\,dx}{x^2+1} = \lim_{b\to+\infty} \frac{1}{2}\ln(x^2+1)\Big|_1^b$$

$$= \lim_{b\to+\infty} \left[\frac{1}{2}\ln(b^2+1) - \frac{1}{2}\ln 2\right] = +\infty,$$

i.e., the integral $\int_1^{+\infty} \frac{x\,dx}{x^2+1}$ diverges, and so does the series. ►

Remark. In the integral $\int_1^{+\infty} f(x)\,dx$ the lower limit may be taken arbitrary, e.g., m, where $m \geqslant 1$.

(4) Examine for convergence the series $\sum_{n=4}^{\infty} \frac{1}{(n-2)\ln^2(n-2)}$.

◄ Since the nth term is $a_n = 1/[(n-2)\ln^2(n-2)]$, $f(x)$ will be

$$f(x) = \frac{1}{(x-2)\ln^2(x-2)},$$

where $x \geqslant 4$. Then

$$\int_4^{+\infty} \frac{dx}{(x-2)\ln^2(x-2)} = \lim_{b\to+\infty} \int_4^b \frac{dx}{(x-2)\ln^2(x-2)}$$

$$= \lim_{b\to+\infty} \int_4^b \frac{d[\ln(x-2)]}{\ln^2(x-2)} = \lim_{b\to+\infty} \left[-\frac{1}{\ln(x-2)}\right]\Bigg|_1^b$$

$$= \lim_{b\to+\infty} \left[\frac{1}{\ln 2} - \frac{1}{\ln(b-2)}\right] = \frac{1}{\ln 2}.$$

Since the integral $\int\limits_4^{+\infty} \frac{dx}{(x-2)\ln^2(x-2)}$ converges, so does the original series. ►

If $\sum\limits_{n=1}^{\infty} f(n)$ converges, the method used to prove the test enables us to estimate the error due to replacing the series sum by a partial sum.

◄ Suppose that $f(x)$ obeys the conditions of Theorem 13.9 and that the series $\sum\limits_{n=1}^{\infty} f(n)$ converges to S. It can be shown that in that case the integral $\int\limits_1^{+\infty} f(x)\,dx$ will converge as well. The remainder R_n of the series will then be

$$R_n = S - S_n = \sum_{k=n+1}^{\infty} f(k) \leqslant \sum_{k=n+1}^{\infty} \int\limits_{k-1}^{k} f(x)\,dx = \int\limits_n^{+\infty} f(x)\,dx,$$

which follows from

$$f(k+1) \leqslant \int\limits_k^{k+1} f(x)\,dx,$$

where $k+1$ is replaced by k.

Thus

$$R_n \leqslant \int\limits_n^{+\infty} f(x)\,dx \quad (R_n > 0).$$

Therefore, if we replace S by S_n, the error will not be larger than $\int\limits_n^{+\infty} f(x)\,dx$. ►

Examples. (1) Examine for convergence the series $\sum\limits_{n=1}^{\infty} \frac{n}{(n^2+1)^2}$ and estimate the error if we replace its sum S by the partial sum S_5.

◄ Here $f(x) = \frac{x}{(x^2+1)^2}$ and

$$\int\limits_1^{+\infty} \frac{x\,dx}{(x^2+1)^2} = \lim_{b\to+\infty} \int\limits_1^{b} \frac{x\,dx}{(x^2+1)^2}$$

$$= \lim_{b\to+\infty} \left[-\frac{1}{2(x^2+1)} \right] \Bigg|_1^b = \lim_{b\to+\infty} \left[\frac{1}{4} - \frac{1}{2(b^2+1)} \right] = \frac{1}{4}.$$

By the integral test the series converges. We denote the sum of that series by S and put $S \approx S_5$. Then

$$S \approx S_5 = \frac{1}{4} + \frac{2}{25} + \frac{3}{100} + \frac{4}{289} + \frac{5}{676}$$
$$= 0.25 + 0.08 + 0.03 + 0.013841 + 0.007396 = 0.381237.$$

Estimate the error R_5. We have

$$R_5 \leqslant \int_5^{+\infty} \frac{x\,dx}{(x^2+1)^2} = -\frac{1}{2(x^2+1)}\bigg|_5^{+\infty} = \frac{1}{52} = 0.019231. \blacktriangleright$$

(2) Estimate the nth remainder of the converging series $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ where $p > 1$.

◄ We have

$$R_n \leqslant \int_n^{+\infty} \frac{dx}{x^p} = \int_n^{+\infty} x^{-p}\,dx = \frac{1}{(1-p)x^{p-1}}\bigg|_n^{+\infty} = \frac{1}{(p-1)n^{p-1}}.$$

Thus,

$$R_n \leqslant \frac{1}{(p-1)n^{p-1}}, \quad n = 1, 2, \ldots . \blacktriangleright$$

13.4 Alternating Series. Leibniz Test

Definition. The number series

$$a_1 - a_2 + a_3 - \ldots + (-1)^{n-1} a_n + \ldots,$$

where all a_n are of one sign (e.g., $a_n > 0$), is called the *alternating series* or a *series of terms with alternating signs.*

For example, the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

is alternating, and the series

$$1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} - \frac{1}{6} + \ldots$$

is not.

The following test, known as the *Leibniz test*, holds for alternating series.

Theorem 13.10. *Suppose that in the alternating series* $a_1 - a_2 + a_3 - \dots$ *all* a_n *are such that* $a_1 > a_2 > a_3 > \dots$ *and* $\lim_{n\to\infty} a_n = 0$. *Then, the series converges, its sum* S *is positive and does not exceed the first term, i.e.,* $0 < S \leqslant a_1$.

◄ We take the even partial sum S_{2n} and write it as

$$S_{2n} = (a_1 - a_2) + (a_3 - a_4) + \dots + (a_{2n-1} - a_{2n}).$$

It follows from the statement of the theorem that the differences in parentheses are positive, and so S_{2n} increases with n and $S_{2n} > 0$.

The sum can be written as

$$S_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \dots - (a_{2n-2} - a_{2n-1}) - a_{2n},$$

where each difference is positive. It follows that $S_{2n} < a_1$ ($n = 1, 2, \dots$). The sequence $\{S_{2n}\}$ thus increases monotonically and is bounded, i.e., $0 < S_{2n} < a_1$ for all n. Consequently, it has the limit $\lim_{n\to\infty} S_{2n} = S$ such that $0 < S \leqslant a_1$.

The odd partial sum S_{2n+1} will be

$$S_{2n+1} = S_{2n} + a_{2n+1} \quad (n = 1, 2, \dots).$$

We have proved that $\lim_{n\to\infty} S_{2n} = S$ and by the statement of the theorem $\lim_{n\to\infty} a_{2n+1} = 0$. Therefore, there exists

$$\lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} S_{2n} + \lim_{n\to\infty} a_{2n+1} = S.$$

We have thus shown that $\lim_{n\to\infty} S_n = S$, i.e., the series converges. In particular, from the inequality $0 < S \leqslant a_1$ it follows that the sum of the series is positive. ►

Remark. The theorem is valid if the condition that $\{S_n\}$ be monotonic is met starting with a certain N for all $n \geqslant N$, so that the discarding of a finite number of terms does not affect the convergence of the series.

Example. The alternating series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{(-1)^{n-1}}{n} + \dots$$

converges, since

$$1 > \frac{1}{2} > \frac{1}{3} > \dots \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} = 0.$$

Theorem 13.10 allows us to estimate the nth remainder $R_n = \pm(a_{n+1} - a_{n+2} + \dots)$ of the series, which is itself an alternating series. We have $|R_n| \leqslant a_{n+1}$. Since $R_n = S - S_n$, then

$$|S - S_n| \leqslant a_{n+1}.$$

The absolute error due to the replacement of the sum of an alternating series by its nth partial sum is not larger in absolute value than the first of the discarded terms.

Example. Compute approximately the sum of the series

$$1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \dots + \frac{(-1)^{n-1}}{n!} + \dots$$

keeping only the first four terms, and estimate the error.

◄ The convergence is obvious. We put approximately $S \approx S_4 = 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} = 0.625$. Then $|S - S_4| \leqslant \frac{1}{5!} = \frac{1}{120}$. The absolute error is thus not larger than $1/120 = 0.0083$. ►

13.5 Series of Positive and Negative Terms. Absolute and Conditional Convergence

Definition. The number series $\sum_{n=1}^{\infty} a_n$ of real terms of any sign is called a *series of positive and negative terms*.

Examples are:

$$1 + \cos 1 + \cos 2 + \dots + \cos n + \dots$$

$$1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} - \frac{1}{6} + \dots$$

(plus, two minuses, plus, two minuses, and so on). Obviously, the series considered in the previous section is a special case of a series of positive and negative terms.

Along with the series of positive and negative terms $a_1 + a_2 + \dots$ we consider the series $|a_1| + |a_2| + \dots$ and prove the following theorem:

Theorem 13.11. *If the series* $\sum_{n=1}^{\infty} |a_n|$ *converges, then the series* $\sum_{n=1}^{\infty} a_n$ *converges as well.*

◄ From the inequality $-|a_n| \leqslant a_n \leqslant |a_n|$ we get $0 \leqslant a_n + |a_n| \leqslant 2|a_n|$ for $n = 1, 2, \dots$. Let the series $\sum_{n=1}^{\infty} |a_n|$ be convergent, then the series

$\sum_{n=1}^{\infty} 2|a_n|$ will be convergent as well, and from the comparison test the series $\sum_{n=1}^{\infty}(a_n + |a_n|)$ will be convergent too. But $\sum_{n=1}^{\infty} a_n$ is the difference of two convergent series

$$\sum_{n=1}^{\infty}(a_n + |a_n|) - \sum_{n=1}^{\infty}|a_n| = \sum_{n=1}^{\infty} a_n,$$

therefore it will converge. ►

Corollary. If $\sum_{n=1}^{\infty}|a_n|$ converges, then we have $\left|\sum_{n=1}^{\infty} a_n\right| \leqslant \sum_{n=1}^{\infty}|a_n|$.

◄ For any natural k we have

$$\left|\sum_{n=1}^{k} a_n\right| \leqslant \sum_{n=1}^{k}|a_n|,$$

i.e.,

$$-\sum_{n=1}^{k}|a_n| \leqslant \sum_{n=1}^{k} a_n \leqslant \sum_{n=1}^{k}|a_n|.$$

Passing to the limit as $k \to \infty$ gives

$$-\sum_{n=1}^{\infty}|a_n| \leqslant \sum_{n=1}^{\infty} a_n \leqslant \sum_{n=1}^{\infty}|a_n|$$

or

$$\left|\sum_{n=1}^{\infty} a_n\right| \leqslant \sum_{n=1}^{\infty}|a_n|. \text{ ►}$$

When examining $\sum_{n=1}^{\infty}|a_n|$ for convergence we can make use of all sufficient tests established for series of positive terms.

Remark. Generally speaking, the convergence of $\sum_{n=1}^{\infty} a_n$ does not suggest the convergence of $\sum_{n=1}^{\infty}|a_n|$, i.e., the theorem only gives a sufficient condition for the convergence of $\sum_{n=1}^{\infty} a_n$. In other words, this condition is not a necessary one.

Example. The series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$ converges by the Leibniz test, but the series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$ diverges.

Definitions. The series of positive and negative terms $\sum_{n=1}^{\infty} a_n$ is called *absolutely convergent*, if the series $\sum_{n=1}^{\infty} |a_n|$ converges.

The series $\sum_{n=1}^{\infty} a_n$ is called *conditionally convergent* if it converges and the series $\sum_{n=1}^{\infty} |a_n|$ diverges.

Examples. (1) The number series

$$1 - \frac{1}{2^2} - \frac{1}{3^2} + \frac{1}{4^2} - ...$$

(plus, two minuses, plus, two minuses, and so on) is absolutely convergent, since the series of the absolute values of its terms $1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + ...$ converges.

(2) The number series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...$ is conditionally convergent, since the series of the absolute values of its terms is the harmonic series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + ...$, which diverges.

The absolutely and conditionally convergent series have the following *properties*:

Theorem 13.12. *If the terms of an absolutely convergent series are arbitrarily rearranged, the series remains absolutely convergent, and its sum remains the same.*

A conditionally convergent series does not possess this property.

Theorem 13.13. *If a series is conditionally convergent, then for any specified number A we can rearrange the terms so that the resultant series will have the sum A.*

Moreover, *the terms of a conditionally convergent series can be rearranged so that the resultant series will be divergent.*

By way of example, consider the conditionally convergent series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + ...$ with sum S.

We rearrange the terms so that each positive term be followed by two subsequent negative ones. We will get

$$\left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \left(\frac{1}{5} - \frac{1}{10} - \frac{1}{12}\right) + ...$$
$$= \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{6} - \frac{1}{8}\right) + \left(\frac{1}{10} - \frac{1}{12}\right) + ...$$

$$= \frac{1}{2}\left(1 - \frac{1}{2}\right) + \frac{1}{2}\left(\frac{1}{3} - \frac{1}{4}\right) + \frac{1}{2}\left(\frac{1}{5} - \frac{1}{6}\right) + \ldots$$

$$= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots\right) = \frac{1}{2}S.$$

The rearrangement has thus given us a series with half the sum of the original series.

Exercises

Write the nth term of the series below:

1. $\frac{1}{2} + \frac{8}{3} + \frac{27}{4} + \ldots$. 2. $\frac{1}{2} + \frac{4}{4} + \frac{9}{8} + \frac{16}{16} + \ldots$.

3. $2 + \frac{3}{4} + \frac{4}{9} + \frac{5}{16} + \ldots$. 4. $\frac{1}{2} + \frac{2}{5} + \frac{3}{10} + \frac{4}{17} + \ldots$.

5. $\frac{2}{3} + \frac{6}{8} + \frac{24}{15} + \frac{120}{24} + \ldots$.

Find the sum S_n of the first n terms of the following series and prove its convergence using the definition of the convergence of a series:

6. $\sum_{n=1}^{\infty} \frac{1}{n(n-1)}$. 7. $\sum_{n=1}^{\infty} \frac{1}{9n^2 - 3n - 2}$. 8. $\sum_{n=1}^{\infty} \frac{2^{n+1}}{3^n}$. 9. $\sum_{n=1}^{\infty} \frac{2n+1}{n^2(n+1)^2}$.

Use the comparison test and other relevant tests to examine the following series for convergence:

10. $\sum_{n=1}^{\infty} \frac{2n+1}{n+1}$. 11. $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$. 12. $\sum_{n=1}^{\infty} \frac{\sin^2 \alpha n}{2^n}$ $(\alpha \neq 0)$.

13. $\sum_{n=1}^{\infty} \frac{1}{\sqrt{4^n + 7}}$. 14. $\sum_{n=1}^{\infty} e^{-\sin \frac{\pi}{n}}$ 15. $\sum_{n=1}^{\infty} \frac{1}{2^n + \cos^2 n}$.

16. $\sum_{n=1}^{\infty} \left(1 - \cos \frac{\pi}{3^n}\right)$. 17. $\sum_{n=3}^{\infty} \tan \frac{\pi}{n}$. 18. $\sum_{n=1}^{\infty} \ln(1 + 2^{-n})$.

19. $\sum_{n=1}^{\infty} \sin \frac{\pi}{2\sqrt{n}}$.

Using the D'Alembert test, examine for convergence the series:

20. $\sum_{n=1}^{\infty} \frac{n^3}{3^n}$. 21. $\sum_{n=1}^{\infty} \frac{2^n}{n^2+n}$. 22. $\sum_{n=1}^{\infty} \frac{n^4}{4^n+1}$. 23. $\sum_{n=1}^{\infty} \frac{n^n}{n!}$.

24. $\sum_{n=1}^{\infty} \frac{n!}{n^2 \cdot 2^n}$. 25. $\sum_{n=1}^{\infty} n^3 \sin \frac{\pi}{3^n}$. 26. $\sum_{n=1}^{\infty} n \tan \frac{\pi}{2^{n+1}}$.

27. $\sum_{n=1}^{\infty} \frac{a^n}{n^n} \quad (0 < a \neq 1)$.

Using the Cauchy test, examine for convergence the series:

28. $\sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{n^2}$. 29. $\sum_{n=1}^{\infty} \left(\frac{n}{2n+1}\right)^n$. 30. $\sum_{n=1}^{\infty} (\tan^{-1})^n \frac{1}{n}$. 31. $\sum_{n=1}^{\infty} \frac{8^{n+2}}{5^n}$.

32. $\sum_{n=1}^{\infty} 2^{-n} \left(\frac{n+1}{n}\right)^{n^2}$. 33. $\sum_{n=1}^{\infty} \ln^n \frac{2n+1}{n}$.

Using the integral test, examine for convergence the series:

34. $\sum_{n=1}^{\infty} \frac{n}{\sqrt{n^2+1}}$. 35. $\sum_{n=1}^{\infty} \frac{1}{n^2} \sin \frac{1}{n}$. 36. $\sum_{n=2}^{\infty} \frac{1}{n\sqrt{\ln n}}$. 37. $\sum_{n=1}^{\infty} n e^{-\frac{n^2}{2}}$.

38. $\sum_{n=1}^{\infty} \frac{\tan^{-1} n}{n^2+1}$. 39. $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} 2^{-\sqrt{n}}$.

Combining various tests, examine for convergence the series:

40. $\sum_{n=1}^{\infty} \frac{n^2+n+1}{n^4+n+1}$. 41. $\sum_{n=2}^{\infty} \frac{1}{(n^2+1)\ln n}$.

42. $\sum_{n=2}^{\infty} \frac{1}{(3n+2)\ln^2 n}$. 43. $\sum_{n=1}^{\infty} \frac{n^4}{\ln^4(n+1)}$.

Hint: Use the inequality

$$\ln(1+x) \leqslant x, \quad -1 < x < +\infty.$$

44. $\sum_{n=2}^{\infty} \frac{n}{2^{n^2} + n\ln^2 n}$. 45. $\sum_{n=1}^{\infty} \sin\frac{\pi}{n^2+1}$.

Examine for conditional convergence and absolute convergence the series:

46. $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{n}{2^n}$. 47. $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{\sqrt{n}}$. 48. $\sum_{n=2}^{\infty}\frac{(-1)^{n+1}}{\ln\ln n}$.

49. $\sum_{n=1}^{\infty}\frac{(-1)^{n-1}3^n}{(2n-1)^n}$. 50. $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}n}{6n-5}$. 51. $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{3^n}{n^3}$.

Answers

1. $a_n = \frac{n^3}{n+1}$. 2. $a_n = \frac{n^2}{2^n}$. 3. $a_n = \frac{n+1}{n^2}$. 4. $a_n = \frac{n}{n^2+1}$. 5. $a_n = \frac{n!}{n^2-1}$. 6. $S_n = \frac{n}{n+1}$.
7. $S_n = \frac{n}{3n+1}$. 8. $S_n = 6\left[1 - \left(\frac{2}{3}\right)^{n+1}\right]$. 9. $S_n = \frac{n^2+2n}{(n+1)^2}$

10. Diverges. 11. Diverges. 12. Converges. 13. Converges. 14. Diverges. 15. Converges. 16. Converges. 17. Diverges. 18. Converges. 19. Diverges. 20. Converges. 21. Diverges. 22. Converges. 23. Diverges. 24. Diverges. 25. Converges. 26. Converges. 27. Converges. 28. Converges. 29. Converges. 30. Converges. 31. Diverges. 32. Diverges. 33. Converges. 34. Diverges. 35. Converges. 36. Diverges. 37. Converges. 38. Converges. 39. Converges. 40. Converges. 41. Converges. 42. Converges. 43. Diverges. 44. Converges. 45. Converges. 46. Absolutely converges. 47. Conditionally converges. 48. Conditionally converges. 49. Absolutely converges. 50. Diverges. 51. Diverges.

Chapter 14

Functional Series

14.1 Convergence Domain and Convergence Interval

The series

$$f_1(x) + f_2(x) + \dots + f_n(x) + \dots = \sum_{n=1}^{\infty} f_n(x) \tag{14.1}$$

whose terms are functions $f_n(x)$, $n = 1, 2, \dots$, defined on a certain set E of the number axis, is called the *functional series.*

For example, the terms of the series $1 + x + x^2 + \dots$ are defined on the interval $-\infty < x < +\infty$ and the terms of the series $1 + \sin^{-1}x + (\sin^{-1})^2 x + \dots$ are defined on the interval $-1 \leqslant x \leqslant 1$.

The functional series (14.1) is said to be *convergent* at a point $x_0 \in E$, if the number series $\sum_{n=1}^{\infty} f_n(x_0)$ is convergent. If series (14.1) converges at each point x of a set $D \subseteq E$, (14.1) is said to be convergent on D, and D is called the *convergence domain* of the series.

Series (14.1) is said to be *absolutely convergent* on the set D, if on this set the series $\sum_{n=1}^{\infty} |f_n(x)|$ converges.

If (14.1) converges on D, its sum S is a function of x defined on D:

$$S = S(x), \quad x \in D.$$

For some functional series we can find their convergence intervals using the sufficient tests established for series of positive terms, e.g., D'Alembert's and Cauchy's.

Examples. (1) Find the convergence interval of the series $\sum_{n=1}^{\infty} \frac{1}{n^{\log x}}$.

◄ The number series $\sum_{n=1}^{\infty} 1/n^p$ converges for $p > 1$ and diverges for $p \leqslant 1$, and so if we put $p = \log x$, we obtain the given series, which will converge for $\log x > 1$, i.e., for $x > 10$, and diverge for $\log x \leqslant 1$, i.e., for $0 < x \leqslant 10$. The convergence interval is thus $10 < x < +\infty$. ►

(2) Find the convergence interval for the series $\sum_{n=1}^{\infty}(-1)^{n-1} ne^{nx}$.

◄ Consider the series

$$\sum_{n=1}^{\infty}\left|(-1)^{n-1} ne^{nx}\right| = \sum_{n=1}^{\infty} ne^{nx}. \tag{14.2}$$

Since the terms are positive for all x, we will apply, say, the D'Alembert test

$$\lambda = \lim_{n\to\infty}\frac{(n+1)e^{(n+1)x}}{ne^{nx}} = \lim_{n\to\infty}\frac{(n+1)e^{x}}{n} = e^{x}.$$

Series (14.2) will converge if $e^x < 1$, i.e., for $x < 0$. Consequently, the series is absolutely convergent in the interval $-\infty < x < 0$. When $x > 0$, the series diverges, since $\lambda = e^x > 1$; it is obvious that at $x = 0$ the series diverges. ►

(3) Find the convergence interval for the series $\sum_{n=1}^{\infty}\frac{n^n}{(1+x^2)^n}$.

◄ The terms are defined, continuous and positive in the interval $-\infty < x < +\infty$. Applying the Cauchy test, we obtain

$$\lambda = \lim_{n\to\infty}\sqrt[n]{\frac{n^n}{(1+x^2)^n}} = \lim_{n\to\infty}\frac{n}{1+x^2} = +\infty$$

for all $x \in (-\infty, +\infty)$. The series is thus divergent for all x. ►

Let $S_n(x)$ be the nth partial sum of the functional series $\sum_{n=1}^{\infty} f_n(x)$. If the series converges on the set D and its sum is $S(x)$, then it can be represented as

$$S(x) = S_n(x) + R_n(x),$$

where $R_n(x)$ is the sum of the convergent (on D) series

$$f_{n+1}(x) + f_{n+2}(x) + \ldots,$$

i.e.,

$$R_n(x) = f_{n+1}(x) + f_{n+2}(x) + \ldots = \sum_{k=n+1}^{\infty} f_k(x).$$

The quantity $R_n(x)$ is called the nth *remainder* of the functional series $\sum_{n=1}^{\infty} f_n(x)$.

Take any (arbitrarily small) number $\varepsilon > 0$. Then if (14.4) converges there exists $N = N(\varepsilon)$ such that $\sigma - \sigma_n < \varepsilon$, and hence $|S(x) - S_n(x)| < \varepsilon$ for all $n > N(\varepsilon)$ and for all $x \in \Omega$, i.e., series (14.3) converges uniformly on Ω. ►

Remark. Series (14.4) is often called the *dominant* series for the functional series (14.3).

Examples. (1) Examine for uniform convergence the series $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$.

◄ The inequality

$$\left| \frac{\cos nx}{n^2} \right| = \frac{|\cos nx|}{n^2} \leqslant \frac{1}{n^2}$$

holds for all $n = 1, 2, \ldots$ and for all $x \in (-\infty, +\infty)$. The number series $\sum_{n=1}^{\infty} 1/n^2$ converges and by the Weierstrass test the original series converges uniformly and absolutely on the entire axis $-\infty < x < +\infty$. ►

(2) Examine for uniform convergence the series $\sum_{n=1}^{\infty} \frac{\sin nx}{n^2 + (4 - x^2)^{n/2}}$.

◄ The terms are defined and continuous on the interval $[-2, 2]$. Since $(4 - x^2)^{n/2} = (\sqrt{4 - x^2})^n \geqslant 0$ on the interval $[-2, 2]$ for any natural n, we obtain

$$\left| \frac{\sin nx}{n^2 + (4 - x^2)^{n/2}} \right| = \frac{|\sin nx|}{n^2 + (4 - x^2)^{n/2}} \leqslant \frac{1}{n^2 + (4 - x^2)^{n/2}} \leqslant \frac{1}{n^2}.$$

So, the inequality

$$\left| \frac{\sin nx}{n^2 + (4 - x^2)^{n/2}} \right| \leqslant \frac{1}{n^2}$$

holds for $n = 1, 2, \ldots$ and for all $x \in [-2, 2]$. Since the number series $\sum_{n=1}^{\infty} 1/n^2$ converges, then by the Weierstrass test the original series converges absolutely and uniformly on the interval $[-2, 2]$. ►

Remark. The functional series (14.3) can also converge uniformly on Ω when there is no dominant number series (14.4), i.e., the Weierstrass test is only a sufficient test for uniform convergence, but not a necessary one.

For example, as it has been shown above, the series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n + \sqrt{1 - x^2}}$

converges uniformly on the interval $[-1, 1]$, but for it there is no dominant convergent number series (14.4). In fact, the inequality

$$\left|\frac{(-1)^{n-1}}{n+\sqrt{1-x^2}}\right| \leqslant \frac{1}{n}$$

holds for all natural n and for all $x \in [-1, 1]$ (the equality occurs at $x = -1$ and $x = 1$), and the number series $\sum_{n=1}^{\infty} 1/n$ diverges.

14.4 Properties of Uniformly Convergent Functional Series

Some important properties of the uniformly convergent functional series may be given by the following theorems.

Theorem 14.2. *If all the terms of the series* $\sum_{n=1}^{\infty} f_n(x)$ *that is uniformly convergent on the interval* $[a, b]$ *are multiplied by the same function* $g(x)$ *that is bounded on* $[a, b]$, *then the resultant functional series* $\sum_{n=1}^{\infty} g(x) f_n(x)$ *uniformly converges on* $[a, b]$.

◄ Suppose that on $[a, b]$ the series $\sum_{n=1}^{\infty} f_n(x)$ uniformly converges to $S(x)$. Since $g(x)$ is bounded, there exists a constant $C > 0$ such that $|g(x)| \leqslant C \ \forall x \in [a, b]$.

By the definition of uniform convergence, for any number $\varepsilon > 0$ there exists a number N such that for all $n > N$ and for all $x \in [a, b]$ there holds the inequality

$$|S_n(x) - S(x)| < \frac{\varepsilon}{C},$$

where $S_n(x)$ is a partial sum of the original series. Therefore, we will have

$$|g(x)S_n(x) - g(x)\,S(x)| = |g(x)|\;|S_n(x) - S(x)| < C \cdot \frac{\varepsilon}{C} = \varepsilon$$

for $n > N$ and for any $x \in [a, b]$, i.e., $\sum_{n=1}^{\infty} g(x) f_n(x)$ uniformly converges on $[a, b]$ to $g(x)\,S(x)$. ►

Theorem 14.3. *If the series*

$$S(x) = \sum_{n=1}^{\infty} f_n(x)$$

converges uniformly on the interval $[a, b]$ and all its terms are continuous, then its sum $S(x)$ is also continuous on that interval.

◄ We take two arbitrary points x and $x + \Delta x$ in the interval $[a, b]$. Since the given series converges uniformly on $[a, b]$, there will be a number $N = N(\varepsilon)$ for any number $\varepsilon > 0$ such that for all $n > N$

$$|S(x) - S_n(x)| < \frac{\varepsilon}{3} \tag{14.6}$$

and

$$|S(x + \Delta x) - S_n(x + \Delta x)| < \frac{\varepsilon}{3}, \tag{14.7}$$

where $S_n(x)$ are the partial sums of the series $\sum_{n=1}^{\infty} f_n(x)$. The partial sums $S_n(x)$ are continuous on the interval $[a, b]$ as are the sums of a finite number of functions $f_n(x)$. Therefore, there will be a number $\delta = \delta(\varepsilon) > 0$ for a fixed $n_0 > N(\varepsilon)$ and given ε such that for Δx satisfying $|\Delta x| < \delta$ we will have

$$|S_{n_0}(x + \Delta x) - S_{n_0}(x)| < \frac{\varepsilon}{3}. \tag{14.8}$$

The increment ΔS of $S(x)$ can be written as

$$\Delta S = S(x + \Delta x) - S(x) = [S(x + \Delta x) - S_{n_0}(x + \Delta x)]$$
$$+ [S_{n_0}(x + \Delta x) - S_{n_0}(x)] + [S_{n_0}(x) - S(x)],$$

whence

$$|\Delta S| \leqslant |S(x + \Delta x) - S_{n_0}(x + \Delta x)|$$
$$+ |S_{n_0}(x + \Delta x) - S_{n_0}(x)| + |S(x) - S_{n_0}(x)|.$$

From (14.6-8) we obtain the following inequality for Δx satisfying $|\Delta x| < \delta$:

$$|\Delta S| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

This means that $\lim_{\Delta x \to \infty} \Delta S = 0$, i.e., the sum $S(x)$ is continuous at the point x. Since x is an arbitrary point in $[a, b]$, $S(x)$ must be continuous on $[a, b]$. ►

Remark. The functional series $\sum_{n=1}^{\infty} f_n(x)$ whose terms are continuous on $[a, b]$ but which does not uniformly converge on $[a, b]$ may have as its sum a discontinuous function.

Examples. (1) Consider the functional series

$$(1 - x) + x(1 - x) + x^2(1 - x) + \dots + x^{n-1}(1 - x) + \dots$$

on $[0, 1]$. We take its nth partial sum

$$S_n(x) = (1 - x) + x(1 - x) + x^2(1 - x) + \dots + x^{n-1}(1 - x)$$

$$= (1 - x)(1 + x + x^2 + \dots + x^{n-1}) = (1 - x)\frac{1 - x^n}{1 - x} = 1 - x^n.$$

Therefore

$$\lim_{n\to\infty} S_n(x) = \begin{cases} 1 & \text{for } 0 \leqslant x < 1, \\ 0 & \text{for } x = 1, \end{cases}$$

i.e., the sum is

$$S(x) = \begin{cases} 1 & \text{for } 0 \leqslant x < 1, \\ 0 & \text{for } x = 1. \end{cases}$$

It is discontinuous on the interval $[0, 1]$, although the terms are continuous on the interval. By the last theorem the original series does not converge uniformly on the interval $[0, 1]$.

(2) Consider the series $\sum\limits_{n=1}^{\infty} 1/n^x$.

◄ It has been shown (Sec. 13.3) that this series converges for $x > 1$. For $x \geqslant 1 + \alpha$, where $\alpha > 0$, the series will converge uniformly, since

$$\frac{1}{n^x} \leqslant \frac{1}{n^{1+\alpha}}$$

and the number series $\sum\limits_{n=1}^{\infty} 1/n^{1+\alpha}$ converges by the Weierstrass test. Consequently, for any $x > 1$ the sum is continuous.

The function

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x}, \quad x > 1$$

is called the *zeta-function* (Riemann function) and plays a significant role in the theory of numbers.

Theorem 14.4. *If the series*

$$S(x) = \sum_{n=1}^{\infty} f_n(x)$$

converges uniformly on the interval $[a, b]$ *and all its terms are continuous, then*

$$\int_{x_0}^{x} S(t)\,dt = \int_{x_0}^{x} \left[\sum_{n=1}^{\infty} f_n(t)\right] dt = \sum_{n=1}^{\infty} \int_{x_0}^{x} f_n(t)\,dt,$$

i.e., the series can be integrated termwise from x_0 *to* x *for any* x *and* $x_0 \in [a, b]$. *The resultant series will converge uniformly in* x *on the interval* $[a, b]$ *for any* $x_0 \in [a, b]$.

◄ Functions $f_n(x)$ are continuous and the series converges uniformly on $[a, b]$, therefore its sum $S(x)$ is continuous, and hence integrable, on $[a, b]$. Consider the difference

$$\int_{x_0}^{x} S(t)\,dt - \int_{x_0}^{x} S_n(t)\,dt = \int_{x_0}^{x} [S(t) - S_n(t)]\,dt,$$

where x, $x_0 \in [a, b]$.

Since the series converges uniformly on $[a, b]$, for any $\varepsilon > 0$ there is $N(\varepsilon) > 0$ such that for all $n > N(\varepsilon)$ and all $x \in [a, b]$ we will have

$$|S(x) - S_n(x)| < \frac{\varepsilon}{b - a}.$$

But then

$$\left| \int_{x_0}^{x} S(t)\,dt - \int_{x_0}^{x} S_n(t)\,dt \right|$$

$$\leqslant \left| \int_{x_0}^{x} |S(t) - S_n(t)|\,dt \right| < \left| \int_{x_0}^{x} \frac{\varepsilon}{b - a}\,dt \right|$$

$$= \frac{\varepsilon}{b - a} |x - x_0| < \frac{\varepsilon}{b - a} |b - a| = \frac{\varepsilon(b - a)}{b - a} = \varepsilon.$$

Thus,

$$\left| \int_{x_0}^{x} S(t)\,dt - \int_{x_0}^{x} S_n(t)\,dt \right| < \varepsilon$$

for any $n > N(\varepsilon)$. In other words,

$$\int_{x_0}^{x} S(t)\,dt = \lim_{n\to\infty} \int_{x_0}^{x} S_n(t)\,dt$$

$$= \lim_{n\to\infty} \int_{x_0}^{x} \left[\sum_{k=1}^{n} f_k(t)\right] dt = \lim_{n\to\infty} \sum_{k=1}^{n} \int_{x_0}^{x} f_k(t)\,dt,$$

i.e.,

$$\int_{x_0}^{x} S(t)\,dt = \sum_{k=1}^{\infty} \int_{x_0}^{x} f_k(t)\,dt. \blacktriangleright$$

If the series $\sum_{n=1}^{\infty} f_n(t)$ is not uniformly convergent, then, generally speaking, it cannot be integrated term by term, i.e.,

$$\int_{x_0}^{x} S(t)\,dt \neq \sum_{n=1}^{\infty} \int_{x_0}^{x} f_n(t)\,dt.$$

Theorem 14.5. *Suppose that all the terms of the convergent series* $\sum_{n=1}^{\infty} f_n(x)$ *have continuous derivatives and that the series* $\sum_{n=1}^{\infty} f_n'(x)$ *of these derivatives converges uniformly on the interval* $[a, b]$. *Then, at any point* $x \in [a, b]$

$$\left[\sum_{n=1}^{\infty} f_n(x)\right]' = \sum_{n=1}^{\infty} f_n'(x),$$

i.e., the original series can be differentiated termwise.

◄ Denote

$$\sum_{n=1}^{\infty} f_n(x) = S(x), \qquad \sum_{n=1}^{\infty} f_n'(x) = \sigma(x).$$

Take any two points x and $x_0 \in [a, b]$. Then, by virtue of Theorem 14.4, we will have

$$\int_{x_0}^{x} \sigma(t)\,dt = \int_{x_0}^{x} \left[\sum_{n=1}^{\infty} f'(t)\right] dt = \sum_{n=1}^{\infty} \int_{x_0}^{x} f_n'(t)\,dt$$

$$= \sum_{n=1}^{\infty} [f_n(x) - f_n(x_0)] = \sum_{n=1}^{\infty} f_n(x) - \sum_{n=1}^{\infty} f_n(x_0) = S(x) - S(x_0).$$

But since the function $\sigma(x)$ is continuous as the sum of a uniformly convergent series of continuous functions, then by differentiating

$$\int_{x_0}^{x} \sigma(t)\,dt = S(x) - S(x_0)$$

we will get

$$\left[\int_{x_0}^{x} \sigma(t)\,dt\right]' = S'(t), \quad \text{i.e.,} \quad \sigma(x) = S'(x)$$

or

$$\left[\sum_{n=1}^{\infty} f_n(x)\right]' = \sum_{n=1}^{\infty} f_n'(t). \blacktriangleright$$

Exercises

Find the convergence intervals for the functional series below:

1. $\sum_{n=1}^{\infty} \frac{n}{x^n}$. 2. $\sum_{n=1}^{\infty} \frac{\sin nx}{n^2}$. 3. $\sum_{n=1}^{\infty} e^{nx}$. 4. $\sum_{n=1}^{\infty} \frac{n}{2^{nx}}$. 5. $\sum_{n=1}^{\infty} \ln^n x$.

6. $\sum_{n=1}^{\infty} \ln^n (1 + x^2)$. 7. $\sum_{n=1}^{\infty} \frac{n!}{x^n}$. 8. $\sum_{n=1}^{\infty} x^n \tan \frac{x}{2^n}$. 9. $\sum_{n=1}^{\infty} n^{\ln x}$.

10. $\sum_{n=0}^{\infty} \sin \frac{x}{2^n}$.

Using the Weierstrass test, prove that the following functional series converge uniformly in the specified intervals:

11. $\sum_{n=1}^{\infty} \frac{\sin nx}{n^2}$, $-\infty < x < +\infty$. 12. $\sum_{n=1}^{\infty} \frac{\cos nx}{n\sqrt{n}}$, $-\infty < x < +\infty$.

13. $\sum_{n=2}^{\infty} \frac{1}{n \ln^2 n + (4 - x^2)^{n/2}}$, $-2 \leqslant x \leqslant 2$. 14. $\sum_{n=2}^{\infty} \ln\left(1 + \frac{x}{n^3}\right)$, $-1 \leqslant x \leqslant 1$.

Answers

1. $-1 < x < 1$. 2. $-\infty < x < +\infty$. 3. $-\infty < x < 0$. 4. $0 < x < +\infty$. 5. $e^{-1} < x < e$. 6. $-\sqrt{e-1} < x < \sqrt{e-1}$. 7. Diverges everywhere. 8. $-2 < x < 2$. 9. $0 < x < e$. 10. $-\infty < x < +\infty$.

Chapter 15

Power Series

15.1 Abel's Theorem. Interval and Radius of Convergence for Power Series

A series of the type

$$c_0 + c_1x + c_2x^2 + \ldots + c_nx^n + \ldots = \sum_{n=0}^{\infty} c_nx^n \tag{15.1}$$

or

$$c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + c_n(x - x_0)^n + \ldots$$
$$= \sum_{n=0}^{\infty} c_n(x - x_0)^n, \tag{15.2}$$

where $c_0, c_1, c_2, \ldots, c_n, \ldots$ are constant coefficients, is called a *power series* in x. Series (15.1) is obtained from (15.2) by the change $x - x_0 = \tilde{x}$. The power series (15.1) always converges at $x = 0$, and (15.2) at x_0, their sum at these points being c_0.

Example. The series $x + x^3 + \ldots + x^{2n-1} + \ldots$ and $(x + 2)^2 + (x + 2)^4 + \ldots$ are power series.

Now we look at the *convergence intervals* of the power series.

Theorem 15.1 (Abel's theorem). *If a power series* $\sum_{n=0}^{\infty} c_nx^n$ *converges for* $x = x_1 \neq 0$, *then it converges absolutely for all* x, *such that* $|x| < |x_1|$; *if the power series diverges for* $x = x_2$, *then it diverges for any* x *for which* $|x| > |x_2|$.

◄ Suppose that $\sum_{n=1}^{\infty} c_nx^n$ converges at $x = x_1 \neq 0$, i.e., the number series $\sum_{n=0}^{\infty} c_nx_1^n$ converges. Then $\lim_{n\to\infty} c_nx_1^n = 0$, and hence there exists a number $M > 0$ such that $|c_nx_1^n| < M$ for all n. Consider the series $\sum_{n=0}^{\infty} |c_nx^n|$, where $|x| < |x_1|$, and estimate its nth term. We have

$$|c_nx^n| = |c_nx_1^n| \cdot \left|\frac{x}{x_1}\right|^n \leqslant Mq^n,$$

where $q = |x/x_1| < 1$. But $\sum_{n=0}^{\infty} Mq^n$ is a geometric progression with the ratio q, where $0 \leqslant q < 1$, and hence it is a convergent series. By the comparison test, $\sum_{n=0}^{\infty} |c_n x^n|$ converges at any point x for which $|x| < |x_1|$. Consequently, the power series $\sum_{n=0}^{\infty} c_n x^n$ converges absolutely for $|x| < |x_1|$.

Suppose now that $\sum_{n=0}^{\infty} c_n x^n$ diverges for $x = x_2$. Suppose further that the series converges for $|x| > |x_2|$. According to the treatment above it must converge at $x = x_2$, since $|x_2| < |x|$, which is at variance with the divergence condition for $x = x_2$ for the series. ►

Abel's theorem enables convergence intervals to be established for the power series $\sum_{n=0}^{\infty} c_n x^n$. Let the series be convergent at $x_1 \neq 0$. It will then converge absolutely in the interval $(-|x_1|, |x_1|)$. If the series diverges at x_2 (here $|x_2| > |x_1|$), then it will diverge in the infinite intervals $(-\infty, -|x_2|)$ and $(|x_2|, +\infty)$ as well.

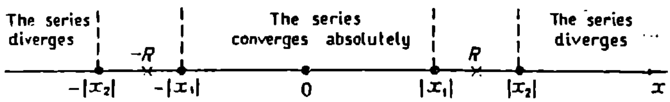


Fig. 15.1

It follows from the above that two points exist on the x-axis (symmetrically about O), which demarcate the interval of convergence from that of divergence (Fig. 15.1).

Theorem 15.2. *There is a unique number $R > 0$ for every power series $\sum_{n=0}^{\infty} c_n x^n$ which converges not only at the point $x = 0$, R being such that the series converges absolutely when $|x| < R$ and diverges when $|x| > R$.*

◄ Let $\mathscr{E}$ be the set of all points x at which the series converges. The set $\mathscr{E}$ is bounded. The theorem states that there are points on the x-axis at which the series diverges. We take one such point, say x_1. Abel's theorem states that for any $x \in \mathscr{E}$ we have $|x| < |x_1|$. However, in a set bounded above there is a unique upper boundary $\sup_{x \in \mathscr{E}} |x|$. Suppose $\sup |x| = R$. Since by definition the series converges not only at $x = 0$, we find that $R > 0$.

We now take any x for which $|x| < R$. By definition of the upper boundary, we can find $x_0 \in \mathscr{E}$ such that $|x| < |x_0| \leqslant R$, whence as follows

from Abel's theorem the series must converge absolutely for the chosen x. If any x is taken for which $|x| > R$, then $x \notin \mathscr{S}$. Consequently, the series must diverge at this x. ►

Thus, the region of (absolute) convergence of a power series $\sum_{n=0}^{\infty} c_n x^n$ is the interval $(-R, R)$ centered on the origin.

Definition. The interval $(-R, R)$, where $R > 0$, at every point $x \in (-R, R)$ of which the series converges and at points such that $|x| > R$ the series diverges, is called the *convergence interval* for the power series $\sum_{n=0}^{\infty} c_n x^n$. The number R is called the *radius of convergence* of the series.

At the ends of the interval $(-R, R)$, i.e., at the points $x = -R$ and $x = R$ the power series either converges or diverges.

Remark. The power series $\sum_{n=0}^{\infty} c_n(x - x_0)^n$ where $x_0 \neq 0$ has the same radius of convergence as $\sum_{n=0}^{\infty} c_n x^n$, but its convergence interval is $(x_0 - R, x_0 + R)$.

When a finite limit

$$\lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|} = L, \quad \text{where } 0 < L < +\infty,$$

exists, the radius of convergence of the power series $\sum_{n=0}^{\infty} c_n x^n$ $\left(\text{or the series } \sum_{n=0}^{\infty} c_n(x - x_0)^n,\ x_0 \neq 0\right)$ can be found from

$$R = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|}. \tag{15.3}$$

To prove (15.3) consider the series composed of the absolute values of the terms of the given one:

$$|c_0| + |c_1 x| + |c_2 x^2| + \ldots + |c_n x^n| + \ldots = \sum_{n=0}^{\infty} |c_n x^n|. \tag{15.4}$$

By applying d'Alembert's test, we find

$$\lim_{n\to\infty} \frac{|c_{n+1} x^{n+1}|}{|c_n x^n|} = \lim_{n\to\infty} \frac{|c_{n+1}|\,|x|^{n+1}}{|c_n|\,|x|^n}$$

$$= \lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|}\,|x| = |x| \lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|} = |x| L.$$

Whence it follows that (15.4) converges if $|x|L < 1$ and diverges if $|x|L > 1$, i.e., a power series converges absolutely for all x such that $|x| < \frac{1}{L}$ and diverges for all $|x| > \frac{1}{L}$. From the definition of the radius of convergence we find that $R = \frac{1}{L}$, i.e.,

$$R = \frac{1}{\lim\limits_{n\to\infty} \frac{|c_{n+1}|}{|c_n|}} \quad \text{or} \quad R = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|}.$$

The convergence radius may also be found from

$$R = \lim_{n\to\infty} \frac{1}{\sqrt[n]{|c_n|}}, \tag{15.5}$$

if a finite limit $\lim\limits_{n\to\infty} \sqrt[n]{|c_n|} = L$, $0 < L < +\infty$ exists. The last identity may easily be obtained from Cauchy's test.

If a power series $\sum\limits_{n=0}^{\infty} c_n x^n$ only converges for $x = 0$, then its convergence radius R is considered to be zero $\left(\text{this is possible, for example, when } \lim\limits_{n\to\infty} \frac{|c_{n+1}|}{|c_n|} = \infty \text{ or } \lim\limits_{n\to\infty} \sqrt[n]{|c_n|} = \infty\right)$. If, however, the series converges at all points along the number axis, then it is considered that $R = +\infty$ (this occurs, for example, when

$$\lim_{n\to\infty} \frac{|c_{n+1}|}{|c_n|} = 0 \quad \text{or} \quad \lim_{n\to\infty} \sqrt[n]{|c_n|} = 0\Big).$$

The convergence domain of a power series $\sum\limits_{n=0}^{\infty} c_n(x - x_0)^n$ may be either an open $(x_0 - R, x_0 + R)$ or a closed $[x_0 - R, x_0 + R]$ interval, or one of the half-intervals $(x_0 - R, x_0 + R]$ or $[x_0 - R, x_0 + R)$. If $R = +\infty$, then the convergence domain is the interval $(-\infty, +\infty)$.

In order to find the convergence domain of a power series $\sum\limits_{n=0}^{\infty} c_n(x - x_0)^n$, first the convergence radius R must be found (using one of the above formulas, for instance) and then the convergence interval $(x_0 - R, x_0 + R)$ in which the series converges absolutely. Second, the convergence of the series must be investigated at the ends of the convergence interval, i.e., at the points $x = x_0 - R$ and $x = x_0 + R$.

Examples. (1) Find the convergence interval of $\sum_{n=1}^{\infty} (-1)^{n-1} n x^n$.

◄ We first find the radius of convergence R using formula (15.3). Since $c_n = (-1)^{n-1} n$ and $c_{n+1} = (-1)^n (n+1)$, we will have

$$R = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = \lim_{n\to\infty} \frac{|(-1)^{n-1} n|}{|(-1)^n (n+1)|} = \lim_{n\to\infty} \frac{n}{n+1} = 1.$$

The radius is $R = 1$, and so the series converges absolutely in the interval $-1 < x < 1$.

Second, we examine the series for convergence at the ends of the convergence interval. Putting $x = -1$, we obtain the number series

$$\sum_{n=1}^{\infty} (-1)^{n-1} n (-1)^n = \sum_{n=1}^{\infty} (-1)^{2n-1} n = \sum_{n=1}^{\infty} (-n),$$

which is obviously divergent, since it does not meet the necessary test: $\lim_{n\to\infty} (-n) \neq 0$. At $x = 1$ we obtain the number series $\sum_{n=1}^{\infty} (-1)^{n-1} n$, for which $\lim_{n\to\infty} (-1)^{n-1} n$ is nonexistent, and hence the series diverges.

Thus, the series converges in the interval $-1 < x < 1$. ►

(2) Find the convergence interval of the series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n 2^n} (x+2)^n$.

◄ (1) We find the radius of convergence by (15.3):

$$c_n = \frac{(-1)^{n-1}}{n 2^n}, \quad c_{n+1} = \frac{(-1)^n}{(n+1) 2^{n+1}};$$

$$R = \lim_{n\to\infty} \frac{\left|\frac{(-1)^{n-1}}{n 2^n}\right|}{\left|\frac{(-1)^n}{(n+1) 2^{n+1}}\right|} = \lim_{n\to\infty} \frac{(n+1) 2^{n+1}}{n 2^n} = \lim_{n\to\infty} 2\left(1 + \frac{1}{n}\right) = 2.$$

The series converges absolutely in the interval $|x+2| < 2$ or $-2 < x + 2 < 2$. Hence $-4 < x < 0$.

(2) At $x = -4$ we obtain the number series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n 2^n} (-2)^n = \sum_{n=1}^{\infty} \frac{(-1)^{2n-1}}{n} = -\sum_{n=1}^{\infty} \frac{1}{n},$$

which diverges (a harmonic series).

At $x = 0$ we will have the number series $\sum_{n=1}^{\infty} (-1)^{n-1}/n$, which converges conditionally. The series thus converges in the interval $-4 < x \leqslant 0$. ►

(3) Find the convergence interval of the series $\sum_{n=1}^{\infty} \frac{(-1)^n (x-2)^n}{n^n}$.

◄ Since $c_n = (-1)^n/n^n$, we will find the radius of convergence using formula (15.5)

$$R = \lim_{n\to\infty} \frac{1}{\sqrt[n]{|c_n|}} = \lim_{n\to\infty} \frac{1}{\sqrt[n]{\left|\frac{(-1)^n}{n^n}\right|}} = \lim_{n\to\infty} n = +\infty.$$

This means that the series converges at all x, i.e., in the interval $(-\infty, +\infty)$. ►

(4) Find the convergence interval of the series

$$\sum_{n=0}^{\infty} n!x^n, \quad 0! = 1.$$

◄ Since $c_n = n!$, $c_{n+1} = (n+1)! = n!(n+1)$, we will get

$$R = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = \lim_{n\to\infty} \frac{n!}{n!(n+1)} = \lim_{n\to\infty} \frac{1}{n+1} = 0.$$

This means that the series converges only at $x = 0$ and its sum is $S(0) = 1$. The convergence interval of the series is one point $x = 0$. ►

15.2 Properties of Power Series

Uniform convergence of a power series. Continuity of its sum. We will need the following theorems.

Theorem 15.3. *The power series* $\sum_{n=0}^{\infty} c_n x^n$ *converges absolutely and uniformly on any interval* $[-a, a]$, $a > 0$, *belonging to the convergence interval* $(-R, R)$, $R > 0$, *of the series.*

◄ Consider the interval $[-a, a]$, where $0 < a < R$. We will have $|c_n x^n| \leqslant |c_n a^n|$ for all x that meet the condition $|x| \leqslant a$, and for $n = 0, 1, 2, \ldots$. But since the number series $\sum_{n=0}^{\infty} |c_n a^n|$ converges, by the Weierstrass test, the given series converges absolutely and uniformly on the interval $[-a, a]$. ►

Theorem 15.4. *The sum* $S(x) = \sum_{n=0}^{\infty} c_n x^n$ *is continuous at each point* x *in the convergence interval* $(-R, R)$, $R > 0$, *of the series.*

◄ Any point x in the convergence interval $(-R, R)$ can be included in the interval $[-a, a]$, $0 < a < R$, where the series converges uniformly. Since the terms of the series are continuous, its sum $S(x)$ will be continuous on the interval $[-a, a]$, and hence at x. ►

Integration of power series. Power series can be integrated according to the following theorem.

Theorem 15.5. *The power series* $\sum_{n=0}^{\infty} c_n x^n$ *can be integrated term by term in its convergence interval* $(-R, R)$, $R > 0$, *and the radius of convergence of the series obtained by termwise integration will also be* R. *Specifically, there holds*

$$\int_0^x \left(\sum_{n=0}^{\infty} c_n t^n\right) dt = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}$$

for any x *in the interval* $(-R, R)$.

◄ Any point x in the convergence interval $(-R, R)$ can be included in the interval $[-a, a]$, where $0 < a < R$. On this interval the series will converge uniformly and, since its terms are continuous, it can be integrated term by term by Theorem 14.4, e.g., from 0 to x, where $0 < |x| < R$. Then

$$\int_0^x \left(\sum_{n=0}^{\infty} c_n t^n\right) dt = \sum_{n=0}^{\infty} c_n \int_0^x t^n dt = \sum_{n=0}^{\infty} \frac{c_n x^{n+1}}{n+1}, \quad x \in (-R, R).$$

We now find the radius of convergence R' of the series obtained, under the condition that there exists a finite $\lim_{n\to\infty} |c_n|/|c_{n+1}| = R$. The radius of convergence R' will then be

$$R' = \lim_{n\to\infty} \frac{\left|\frac{c_n}{n}\right|}{\left|\frac{c_{n+1}}{n+1}\right|} = \lim_{n\to\infty} \left(\frac{n+2}{n+1} \frac{|c_n|}{|c_{n+1}|}\right)$$

$$= \lim_{n\to\infty} \frac{n+2}{n+1} \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = 1 \cdot R = R.$$

Thus, integration does not change the radius of convergence of the power series. ►

Differentiation of power series. Differentiation of power series obeys the following theorem.

Theorem 15.6. *The power series* $S(x) = \sum_{n=0}^{\infty} c_n x^n$ *can be differentiated term by term at any point* x *in its convergence interval* $(-R, R)$, $R > 0$, *and*

$$S'(x) = \left(\sum_{n=0}^{\infty} c_n x^n\right)' = \sum_{n=1}^{\infty} n c_n x^{n-1}.$$

◀ Let R be the radius of convergence of the series

$$\sum_{n=0}^{\infty} c_n x^n \tag{15.6}$$

and R' be the radius of convergence of the series

$$\sum_{n=0}^{\infty} n c_n x^{n-1}, \tag{15.7}$$

and let there be a finite or infinite limit

$$\lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = R.$$

Then, we will obtain

$$R' = \lim_{n\to\infty} \frac{|nc_n|}{|(n+1)c_{n+1}|}$$
$$= \lim_{n\to\infty}\left(1 + \frac{1}{n+1}\right) \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} = 1\cdot R = R.$$

Thus, $R' = R$. Denote by $\sigma(x)$ the sum of (15.7). The series converges on the interval $(-R, R)$. Series (15.6) and (15.7) converge uniformly on any interval $[-a, a]$, where $0 < a < R$. All the terms of (15.7) will then be continuous; they are derivatives of the corresponding terms of (15.6). By Theorem 14.5, we will then have $\sigma(x) = S'(x)$ in the interval $[-a, a]$, and hence in the interval $(-R, R)$ as well, since $a < R$. ▶

Corollary. A power series $\sum_{n=0}^{\infty} c_n x^n$ may be termwise differentiated any number of times at any point x in its convergence interval $(-R, R)$, i.e., its sum $S(x)$ has derivatives of all orders at each point $x \in (-R, R)$, viz.,

$$S^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\ldots(n-k+1)c_n x^{n-k} \quad (k = 1, 2, \ldots).$$

The convergence radius of this series is equal to the convergence radius of the original series $S(x) = \sum_{n=0}^{\infty} c_n x^n$.

By applying the theorem to the series containing the first-order derivatives

$$S'(x) = c_1 + 2c_2 x + 3c_3 x^2 + \ldots + nc_n x^{n-1} + \ldots = \sum_{n=1}^{\infty} nc_n x^{n-1},$$

and then to the series containing the second-order derivatives, etc., we get the formula for $S^{(k)}(x)$ for any k.

15.3 Taylor's Series

Definition. A function $f(x)$ is said to be expanded into a power series $\sum_{n=0}^{\infty} c_n x^n$ on the interval $(-R, R)$ if the series converges on the interval and its sum is $f(x)$, i.e.,

$$f(x) = \sum_{n=0}^{\infty} c_n x^n, \quad x \in (-R, R) \tag{15.8}$$

the interval being assumed not to degenerate into a point.

We shall first prove that the function cannot have two different expansions into power series of the form (15.8).

Theorem 15.7. *If a function $f(x)$ can be expanded into a power series* (15.8) *on the interval $(-R, R)$, $R > 0$, then this expansion is unique, i.e., the coefficients of the series* (15.8) *are uniquely defined by its sum.*

◄ Let the function $f(x)$ be expandable in the interval $(-R, R)$, $R > 0$, into a power series

$$f(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n + \ldots . \tag{15.9}$$

By differentiating this series n times, which can be done in the interval $(-R, R)$ due to Theorem 15.6, we get

$$f^{(n)}(x) = 1 \times 2 \times 3 \times \ldots \times (n-1)nc_n$$
$$+ 2 \times 3 \times \ldots \times (n-1)n(n+1)c_{n+1}x + \ldots .$$

When $x = 0$ we obtain

$$f^{(n)}(0) = 1 \times 2 \times 3 \times \ldots \times (n-1)nc_n$$

or

$$f^{(n)}(0) = n!c_n, \quad n = 0, 1, 2, \ldots ,$$

whence

$$c_n = \frac{f^{(n)}(0)}{n!} \tag{15.10}$$

given that $f^{(n)}(0) = f(0)$, $0! = 1$.

The coefficients c_n ($n = 0, 1, 2, \ldots$) of the power series in (15.8) are thus uniquely determined by (15.10). ►

Remark. If $f(x)$ is expanded in powers of the difference $x - x_0$, i.e.,

$$f(x) = \sum_{n=0}^{\infty} c_n(x - x_0)^n, \quad x \in (x_0 - R, x_0 + R), \ R > 0,$$

According to Theorem 15.9, sin x can be expanded into a Taylor series in x in the interval $(-\infty, +\infty)$ that converges to it. Since

$$f^{(n)}(0) = \sin\frac{n\pi}{2} = \begin{cases} 0 & \text{for } n = 0, 2, 4, \ldots, \\ (-1)^{\frac{n+1}{2}} & \text{for } n = 1, 3, 5, \ldots \end{cases}$$

we will have

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \ldots$$
$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}. \tag{15.15}$$

The radius of convergence is $R = +\infty$.

(3) $f(x) = \cos x$. In a like manner, we obtain

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots$$
$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \quad R = +\infty. \tag{15.16}$$

(4) $f(x) = (1 + x)^\alpha$, where $x > -1$ and α is any real number. The function obeys

$$(1 + x)f'(x) = \alpha f(x) \tag{15.17}$$

and the condition $f(0) = 1$. We will look for a power series, such that its sum $S(x)$ would meet (15.13) and the condition $S(0) = 1$. Let

$$S(x) = 1 + c_1x + c_2x^2 + c_3x^3 + \ldots + c_nx^n + \ldots. \tag{15.18}$$

Hence

$$S'(x) = c_1 + 2c_2x + 3c_3x^2 + \ldots + nc_nx^{n-1} + \ldots. \tag{15.19}$$

Substituting (15.18) and (15.19) into (15.14) gives

$$(1 + x)(c_1 + 2c_2x + 3c_3x^2 + \ldots + nc_nx^{n-1} + \ldots)$$
$$= \alpha(1 + c_1x + c_2x^2 + c_3x^3 + \ldots + c_nx^n + \ldots)$$

or

$$c_1 + (c_1 + 2c_2)x + (2c_2 + 3c_3)x^2 + \ldots$$
$$+ [nc_n + (n + 1)c_{n+1}]x^n + \ldots$$
$$= \alpha + \alpha c_1x + \alpha c_2x^2 + \ldots + \alpha c_nx^n + \ldots.$$

Equating the coefficients at the same powers of x on either side gives

$$c_1 = \alpha, \quad c_1 + 2c_2 = \alpha c_1,$$
$$2c_2 + 3c_3 = \alpha c_2, \ \ldots, \ nc_n + (n+1)c_{n+1} = \alpha c_n, \ \ldots .$$

Hence

$$c_1 = \alpha, \quad c_2 = \frac{\alpha(\alpha-1)}{2!}, \quad c_3 = \frac{\alpha(\alpha-1)(\alpha-2)}{3!},$$

$$\ldots, \ c_n = \frac{\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)}{n!}, \ \ldots .$$

Substituting these coefficients into (15.19) gives

$$S(x) = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3$$
$$+ \ldots + \frac{\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)}{n!}x^n + \ldots . \qquad (15.20)$$

We find the radius of convergence of (15.20) for the case where α is not a natural number. We have

$$R = \lim_{n\to\infty}\frac{|c_n|}{|c_{n+1}|} = \lim_{n\to\infty}\frac{\left|\dfrac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}\right|}{\left|\dfrac{\alpha(\alpha-1)\ldots(\alpha-n+1)(\alpha-n)}{(n+1)!}\right|}$$

$$= \lim_{n\to\infty}\frac{n+1}{|\alpha-n|} = \lim_{n\to\infty}\frac{1+\dfrac{1}{n}}{\left|\dfrac{\alpha}{n}-1\right|} = 1.$$

Series (15.20) thus converges for $|x| < 1$, i.e., in the interval $(-1, 1)$.

We prove that the sum $S(x)$ of (15.20) is $(1+x)^\alpha$ in the interval $(-1, 1)$. Consider

$$\varphi(x) = \frac{S(x)}{f(x)} = \frac{S(x)}{(1+x)^\alpha}.$$

Its derivative is

$$\varphi'(x) = \frac{(1+x)S'(x) - \alpha S(x)}{(1+x)^{\alpha+1}} = \frac{0}{(1+x)^{\alpha+1}} = 0$$

for $x \in (-1, 1)$, since $S(x)$ obeys (15.7).

It follows that $\varphi(x) = C = \text{const}$ in $(-1, 1)$. But since $\varphi(0) = S(0)/1 = 1$, then $C = \varphi(0) = 1$, and hence $S(x)/(1 + x)^\alpha = 1$, i.e., $S(x) = (1 + x)^\alpha$ or

$$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!} x^2 + \ldots + \frac{\alpha(\alpha - 1) \ldots (\alpha - n + 1)}{n!} x^n + \ldots, \tag{15.21}$$

where $-1 < x < 1$.

The series obtained is called a *binomial series*, and its coefficients are called binomial.

When α is a real number, the function $(1 + x)^\alpha$ is a polynomial of degree n, since $R_n(x) \equiv 0$ for all $n > \alpha$ and all x.

Two more expansions are worth noting. At $\alpha = -1$ we will have

$$\frac{1}{1 + x} = 1 - x + x^2 - x^3 + \ldots, \quad -1 < x < 1. \tag{15.22}$$

We substitute $-x$ for x in this relation. Then

$$\frac{1}{1 - x} = 1 + x + x^2 + x^3 + \ldots, \quad -1 < x < 1. \tag{15.23}$$

(5) $f(x) = \ln(1 + x)$, $x > -1$. To expand the function into a Taylor series in x, we integrate (15.22) from 0 to x, where $x \in (-1, 1)$. We have

$$\int_0^x \frac{dt}{1 + t} = \int_0^x (1 - t + t^2 - t^3 + \ldots (-1)^{n+1} t^{n+1} + \ldots) dt$$

or

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + \frac{(-1)^{n+1} x^n}{n} + \ldots. \tag{15.24}$$

Relation (15.24) holds in the interval $-1 < x < 1$. Substituting $-x$ for x yields

$$\ln(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \ldots - \frac{x^n}{n} - \ldots, \tag{15.25}$$

where $-1 < x < 1$.

We can show that (15.24) also holds at $x = 1$:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots.$$

The following is a list of Maclaurin series expansions for some common functions:

(a) $e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots, \quad -\infty < x < +\infty;$

(b) $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \ldots,$
$-\infty < x < +\infty;$

(c) $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots,$
$-\infty < x < +\infty;$

(d) $$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!} x^2 + \frac{\alpha(\alpha - 1)(\alpha - 2)}{3!} x^3 + \ldots$$
$$+ \frac{\alpha(\alpha - 1)(\alpha - 2) \ldots (\alpha - n + 1)}{n!} x^n + \ldots, \quad -1 < x < 1;$$

$$\frac{1}{1 + x} = 1 - x + x^2 - x^3 + \ldots + (-1)^n x^n + \ldots, \quad -1 < x < 1;$$

(e) $\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^{n-1} \frac{x^n}{n} + \ldots, -1 < x \leqslant 1.$

Using these expressions we can deduce expansions for composite functions. It is not necessary then to examine the remainder $R_n(x)$ in the Taylor formula, since the convergence intervals of the above series have been established previously. We will illustrate this by examples.

Examples. (1) Expand the function $1/(4 - x)$ into a power series about $x_0 = 2$, i.e., in powers of $x - 2$.

◄ We transform the function so that we could use series (15.23) for $1/(1 - x)$. Then,

$$\frac{1}{4 - x} = \frac{1}{4 - [(x - 2) + 2]} = \frac{1}{2 - (x - 2)} = \frac{1}{2} \frac{1}{1 - \left(\frac{x - 2}{2}\right)}.$$

Substituting $\frac{x - 2}{2}$ for x in (15.23) we obtain

$$\frac{1}{4 - x} = \frac{1}{2} \left[1 + \frac{x - 2}{2} + \left(\frac{x - 2}{2}\right)^2 + \left(\frac{x - 2}{2}\right)^3 + \ldots\right]$$

or

$$\frac{1}{4 - x} = \frac{1}{2} + \frac{x - 2}{2^2} + \frac{(x - 2)^2}{2^3} + \frac{(x - 2)^3}{2^4} + \ldots .$$

This expansion is valid when

$$\left|\frac{x - 2}{2}\right| < 1, \quad |x - 2| < 2, \quad -2 < x - 2 < 2, \quad 0 < x < 4. \ ►$$

(2) Expand $1/(x^2 - 3x + 2)$ in powers of x (i.e., about $x_0 = 0$), using the expansion (15.23).

◄ We factor the denominator and represent the given rational function as a difference of two simple fractions:

$$\frac{1}{x^2 - 3x + 2} = \frac{1}{(x-1)(x-2)} = \frac{1}{x-2} - \frac{1}{x-1}.$$

After some algebra we get

$$\frac{1}{x^2 - 3x + 2} = \frac{1}{1-x} - \frac{1}{2}\frac{1}{1-\left(\frac{x}{2}\right)}. \tag{15.26}$$

To each term on the right-hand side of (15.26) we apply (15.23) to get

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots, \tag{15.27}$$

$$\frac{1}{1-\frac{x}{2}} = 1 + \frac{x}{2} + \left(\frac{x}{2}\right)^2 + \left(\frac{x}{2}\right)^3 + \ldots. \tag{15.28}$$

Series (15.27) converges for $|x| < 1$, and (15.28) converges for $|x/2| < 1$, i.e., for $|x| < 2$. Both (15.27) and (15.28) converge simultaneously for $|x| < 1$. Since in the interval $(-1, 1)$ the series (15.27) and (15.28) are convergent, we can subtract them from each other term by term. The desired power series will thus be

$$\frac{1}{x^2 - 3x + 2} = (1 + x + x^2 + x^3 + \ldots)$$

$$- \frac{1}{2}\left[1 + \frac{x}{2} + \left(\frac{x}{2}\right)^2 + \left(\frac{x}{2}\right)^3 + \ldots\right] = \frac{1}{2} + \frac{2^2 - 1}{2^2}x$$

$$+ \frac{2^3 - 1}{2^3}x^2 + \ldots + \frac{2^n - 1}{2^n}x^n + \ldots,$$

which is absolutely convergent for $|x| < 1$ (the radius of convergence of the series is $R = 1$). ►

(3) Expand $\sin^{-1} x$ into a Taylor series about $x_0 = 0$.

◄ We know that

$$(\sin^{-1} x)' = \frac{1}{\sqrt{1 - x^2}} = [1 + (-x^2)]^{-1/2}.$$

We apply to this derivative the formula (15.21) with $-x^2$ substituted for x. As a result, for $|-x^2| = x^2 < 1$, i.e., for $-1 < x < 1$, we get

$$(\sin^{-1} x)' = 1 + \frac{1}{2}x^2 + \frac{-\frac{1}{2}\left(-\frac{1}{2}-1\right)}{2!}x^4$$

$$-\frac{-\frac{1}{2}\left(-\frac{1}{2}-1\right)\left(-\frac{1}{2}-2\right)}{3!}x^6 + \ldots$$

$$= 1 + \frac{1}{2}x^2 + \frac{1\times 3}{2!2^2}x^4 + \frac{1\times 3\times 5}{3!2^3}x^6 + \ldots .$$

Integrating this from 0 to x gives

$$\int_0^x (\sin^{-1} t)'\,dt = \int_0^x \left(1 + \frac{t^2}{2} + \frac{1\times 3}{2!2^2}t^4 + \frac{1\times 3\times 5}{3!2^3}t^6 + \ldots\right)dt$$

or

$$\sin^{-1} t\,\Big|_0^x = t\,\Big|_0^x + \frac{1}{2}\,\frac{t^3}{3}\,\Big|_0^x + \frac{1\times 3}{2!2^2}\,\frac{t^5}{5}\,\Big|_0^x$$

$$+ \frac{1\times 3\times 5}{3!2^3}\,\frac{t^7}{7}\,\Big|_0^x + \ldots,$$

i.e.,

$$\sin^{-1} x = x + \frac{1}{2\times 3}x^3 + \frac{1\times 3}{2!2^2\times 5}x^5 + \frac{1\times 3\times 5}{3!2^3\times 7}x^7 + \ldots,$$

where $-1 < x < 1$, $R = 1$.

Termwise integration here is legitimate, since the power series converges uniformly in any interval with ends at 0 and x that is contained in $(-1, 1)$. ►

Remark. Power series expansions can be utilized to take integrals that cannot be expressed in a finite form in terms of elementary functions.

Consider several examples.

Examples. (1) Take the integral

$$\mathrm{Si}\ x = \int_0^x \frac{\sin t}{t}\,dt.$$

◄ The primitive for $(\sin t)/t$ cannot be expressed here through elementary functions. We expand the integrand into a power series, taking into account the equality

$$\sin t = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \ldots . \tag{15.29}$$

From (15.29) we find

$$\frac{\sin t}{t} = 1 - \frac{t^2}{3!} + \frac{t^4}{5!} - \frac{t^6}{7!} + \dots \qquad (15.30)$$

Note that it is legitimate to divide (15.29) by t at $t \neq 0$. At $t = 0$ we set $(\sin t)/t = 1$, and then (15.30) remains valid. Series (15.30) converges at all t ($R = +\infty$). Integrating it term by term gives

$$\operatorname{Si} x = \int_0^x \frac{\sin t}{t}\, dt = x - \frac{x^3}{3!3} + \frac{x^5}{5!5} - \frac{x^7}{7!7} + \dots$$

The series obtained is alternating, so that it is a straightforward matter to estimate the error due to replacement of its sum by a partial sum. ▶

(2) Take the integral $\int_0^x e^{-t^2} dt$.

◀ Here too the primitive for the integrand e^{-t^2} cannot be expressed in terms of elementary functions. To carry out the integration we replace x by $-t^2$ in

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

We get

$$e^{-t^2} = 1 - t^2 + \frac{t^4}{2!} - \frac{t^6}{3!} + \dots$$

We integrate both sides of this from 0 to x:

$$\int_0^x e^{-t^2} dt = t\Big|_0^x - \frac{t^3}{3}\Big|_0^x + \frac{t^5}{2!5}\Big|_0^x - \frac{t^7}{3!7}\Big|_0^x + \dots$$
$$= x - \frac{x^3}{3} + \frac{x^5}{2!5} - \frac{x^7}{3!7} + \dots$$

The series converges for any x (its radius of convergence is $R = +\infty$) and it is alternating for $x > 0$.

Exercises

Find convergence intervals for the following series:

1. $\sum\limits_{n=1}^{\infty} \frac{x^n}{n2^n}$. 2. $\sum\limits_{n=0}^{\infty} \frac{2^n}{n+1} x^n$. 3. $\sum\limits_{n=0}^{\infty} \frac{n+1}{\sqrt{1+2n}} x^n$. 4. $\sum\limits_{n=1}^{\infty} n^n x^n$.

5. $\sum_{n=1}^{\infty} nx^n$. 6. $\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^n$. 7. $\sum_{n=1}^{\infty} \frac{(-1)^n n}{n^2+1} (x+2)^n$.

8. $\sum_{n=1}^{\infty} \frac{(-1)^n n^3}{3^n} (x+3)^n$. 9. $\sum_{n=1}^{\infty} \frac{n^n}{n!} (x+1)^n$. 10. $\sum_{n=1}^{\infty} \frac{(x-1)^n}{\ln(1+n)}$.

11. $\sum_{n=0}^{\infty} \frac{(n+1)!}{5^n} (x-2)^n$. 12. $\sum_{n=1}^{\infty} (-1)^n \frac{n!}{n^2} (x-1)^n$.

Expand the following functions into Maclaurin's series and indicate convergence intervals for the series obtained.

13. $\frac{1}{2+3x}$. 14. $\frac{x}{x-1}$. 15. $\sinh x$. 16. $\cosh x$. 17. $\sqrt[3]{1+x^3}$.

18. $\frac{x}{\sqrt[3]{8+x}}$. 19. $\tan^{-1} x$. 20. $\ln(x+\sqrt{1+x^2})$. 21. a^x, $0 < a \neq 1$.

22. $\sin\left(x - \frac{\pi}{4}\right)$. 23. $\cos\left(2x - \frac{\pi}{4}\right)$. 24. $\sin^2 x$.

Hint. Use the expansions (a)-(e) on page 67.

Expand the following functions into Taylor's series about the point indicated and find convergence intervals of the series obtained using the expansions (a)-(e).

25. $\frac{1}{x}$, $x_0 = -3$. 26. $\frac{1}{2x+3}$, $x_0 = 1$. 27. $\frac{x+1}{x-1}$, $x_0 = -1$.

28. $\frac{x+1}{x+2}$, $x_0 = +1$. 29. $\sin x$, $x_0 = \frac{\pi}{2}$. 30. $\cos x$, $x_0 = \frac{\pi}{2}$.

31. $\ln(3x-2)$, $x_0 = 2$. 32. $\ln(2x-1)$, $x_0 = 1$. 33. e^{-x}, $x_0 = -1$.

34. $\sqrt{x+1}$, $x_0 = 1$.

Answers

1. $-2 \leqslant x < 2$. 2. $-\frac{1}{2} \leqslant x < \frac{1}{2}$. 3. $-1 < x < 1$. 4. $x = 0$. 5. $-1 < x < 1$. 6. $-1 < x \leqslant 1$. 7. $-3 < x \leqslant -1$. 8. $-6 < x < 0$. 9. $-e^{-1} - 1 < x < e^{-1} - 1$. 10. $-1 < x < 3$. 12. $x = 1$. 13. $\frac{1}{2} - \frac{3}{2^2}x + \frac{3^2}{2^3}x^2 - \ldots$, $-\frac{2}{3} < x < \frac{2}{3}$.

14. $-x - x^2 - x^3 - \dots,\quad -1 < x < 1.$ 15. $x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots,\quad -\infty < x < +\infty.$ 16. $1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots,\quad -\infty < x < +\infty.$ 17. $1 + \frac{1}{3}x^3 - \frac{1 \times 2}{2!\,3^2}x^6 + \frac{1 \times 2 \times 5}{3!\,3^3}x^9 - \dots,\quad -1 < x < 1.$ 18. $\frac{1}{2}x - \frac{1}{3 \times 8 \times 2}x^2 + \frac{1 \times 4}{2!\,3^2 \times 8^2 \times 2}x^3 - \frac{1 \times 4 \times 7}{3!\,3^3 \times 8^3 \times 2}x^4 + \dots,\quad -1 < x < 1.$ 19. $x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7},\quad -1 \leqslant x \leqslant 1.$ 20. $x - \frac{x^3}{2 \times 3} + \frac{1 \times 3}{2 \times 4 \times 5}x^5 - \dots + (-1)^n \frac{1 \times 3 \times 5 \times \dots \times (2n-1)}{2 \times 4 \times 6 \times \dots \times 2n(2n+1)}x^{2n+1} + \dots,\quad -1 \leqslant x \leqslant 1.$ 21. $a^x = 1 + \sum_{n=1}^{\infty} \frac{\ln^n a}{n!}x^n,\quad -\infty < x < +\infty.$ 22. $\frac{\sqrt{2}}{2}\left(-1 + x + \frac{x^2}{2!} - \frac{x^3}{3!} - \frac{x^4}{4!} + \dots\right),\quad -\infty < x < +\infty.$ 23. $\frac{\sqrt{2}}{2}\left(1 + 2x - \frac{2^2}{2!}x^2 - \frac{2^3}{3!}x^3 + \frac{2^4}{4!}x^4 + \frac{2^5}{5!}x^5 - \dots\right),\quad -\infty < x < +\infty.$ 24. $\frac{2}{2!}x^2 - \frac{2^3}{4!}x^4 + \frac{2^5}{6!}x^6 - \dots + (-1)^n \frac{2^{2n-1}}{(2n)!}x^{2n} + \dots,\quad -\infty < x < +\infty.$ 25. $-\frac{1}{3}\left[1 + \frac{x+3}{3} + \frac{(x+3)^2}{3^2} + \frac{(x+3)^3}{3^3} + \dots\right],\quad -6 < x < 0.$ 26. $\frac{1}{2} - \frac{1}{5}(x-1) + \frac{2}{5^2}(x-1)^2 - \frac{2^2}{5^3}(x-1)^3 + \dots,\quad -\frac{3}{2} < x < \frac{7}{2}.$ 27. $-\frac{x+1}{2} - \frac{(x+1)^2}{2^2} - \frac{(x+1)^3}{2^3} - \dots,\quad -3 < x < 1.$ 28. $\frac{2}{3} + \frac{x-1}{3^2} - \frac{(x-1)^2}{3^3} + \frac{(x-1)^3}{3^4} + \dots,\quad -2 < x < 4.$ 29. $1 - \frac{\left(x - \frac{\pi}{2}\right)^2}{2!} + \frac{\left(x - \frac{\pi}{2}\right)^4}{4!} - \frac{\left(x - \frac{\pi}{2}\right)^6}{6!} + \dots,\quad x < +\infty.$ 30. $-\left(x - \frac{\pi}{2}\right) + \frac{\left(x - \frac{\pi}{2}\right)^3}{3!} - \frac{\left(x - \frac{\pi}{2}\right)^5}{5!} + \dots,\quad -\infty < x < +\infty.$ 31. $2\ln 2 + \frac{3}{4}(x-2) - \frac{3^2}{2 \times 4^2}(x-2)^2 + \frac{3^3}{3 \times 4^3}(x-2)^2 - \dots,\quad \frac{2}{3} < x \leqslant \frac{10}{3}.$ 32. $2(x-1) - \frac{2^2(x-1)^2}{2} + \frac{2^3(x-1)^3}{3} - \dots,\quad -\frac{1}{2} < x \leqslant \frac{3}{2}.$ 33. $e\left[1 - (x+1) + \frac{(x+1)^2}{2!} - \frac{(x+1)^3}{3!} + \dots\right],\quad -\infty < x < +\infty.$ 34. $\sqrt{2}\left[1 + \frac{x-1}{2^2} - \frac{1}{2!\,2^4}(x-1)^2 + \frac{1 \times 3}{3!\,2^6}(x-1)^3 - \dots\right],\quad -1 < x < 3.$

Chapter 16

Fourier series

16.1 Trigonometric Series

Definition. A function $f(x)$ defined on an infinite set D is called *periodic* if there exists a number $T \neq 0$ such that for each $x \in D$

$$f(x \pm T) = f(x) \quad (x \pm T \in D).$$

The number T is called the *period* of $f(x)$.

Examples. (1) The function $f(x) = \sin x$ defined in the interval $(-\infty, +\infty)$ is periodic, since there exists $T = 2\pi \neq 0$ such that for all $x \in (-\infty, +\infty)$ we have $\sin(x + 2\pi) = \sin x$, $x + 2\pi \in (-\infty, +\infty)$. Thus, $\sin x$ has the period $T = 2\pi$.

The same applies to $f(x) = \cos x$.

(2) The function $f(x) = \tan x$, defined on the set D of $x \neq \pi/2 + n\pi$ $(n = 0, +1, +2, \ldots)$ is periodic, since there exists the number $T \neq 0$, i.e., $T = \pi$, such that for all $x \in D$ we have $\tan(x + \pi) = \tan x$, where $x + \pi \in D$.

Definition. A functional series of the type

$$\frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x$$

$$+ \ldots + a_n \cos nx + b_n \sin nx + \ldots$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \qquad (16.1)$$

is said to be a *trigonometric series*, and the constants a_0, a_n, b_n $(n = 1, 2, \ldots)$ are called the *coefficients of the trigonometric series* (16.1).

The partial sums $S_n(x)$ of the trigonometric series (16.1) are linear combinations of the functions 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, $\ldots$, which are called a *trigonometric system*. Since the terms of the series are periodic functions with a period of 2π, then, if (16.1) converges, its sum $S(x)$ will be a periodic function with a period of 2π.

$$S(x + 2\pi) = S(x), \quad x \in (-\infty, +\infty).$$

Definition. To expand a periodic function $f(x)$ into a trigonometric series (16.1) means to find a convergent trigonometric series whose sum is equal to $f(x)$, i.e., $f(x) = S(x)$, $x \in (-\infty, +\infty)$.

Orthogonality of trigonometric systems. We will need some basic notions in what follows.

Definitions. Functions $f(x)$ and $g(x)$ that are continuous on the interval $[a, b]$ are called *orthogonal* on the interval, if

$$\int_a^b f(x)g(x)dx = 0.$$

For example, the functions $f(x) = x$ and $g(x) = x^2$ are orthogonal on $[-1, 1]$, since

$$\int_{-1}^{1} x \cdot x^2 dx = \int_{-1}^{1} x^3 dx = 0.$$

A finite or infinite system of functions $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ continuous on the interval $[a, b]$, where $\varphi_n(x) \not\equiv 0$ and $n = 1, 2, \ldots$, is called an *orthogonal system* on the interval $[a, b]$ if for all m and n, such that $m \neq n$, we have

$$\int_b^a \varphi_m(x)\varphi_n(x)dx = 0.$$

Theorem 16.1. *The trigonometric system* $1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots$ *is orthogonal on the interval* $[-\pi, \pi]$.

◄ For any integer $n \neq 0$

$$\int_{-\pi}^{\pi} 1 \cdot \cos nx dx = \frac{\sin nx}{n}\Big|_{-\pi}^{\pi} = 0,$$

$$\int_{-\pi}^{\pi} 1 \cdot \sin nx dx = -\frac{\cos nx}{n}\Big|_{-\pi}^{\pi} = 0.$$

Using the well-known trigonometric formulas

$$\cos\alpha\cos\beta = \frac{\cos(\alpha - \beta) + \cos(\alpha + \beta)}{2},$$

$$\sin\alpha\sin\beta = \frac{\cos(\alpha - \beta) - \cos(\alpha + \beta)}{2},$$

for any natural m and n, $m \neq n$, we find

$$\int_{-\pi}^{\pi} \cos mx \cos nx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} [\cos(m-n)x + \cos(m+n)x]\,dx$$

$$= \frac{1}{2}\left[\frac{\sin(m-n)x}{m-n}\Bigg|_{-\pi}^{\pi} + \frac{\sin(m+n)x}{m+n}\Bigg|_{-\pi}^{\pi}\right] = 0,$$

$$\int_{-\pi}^{\pi} \sin mx \sin nx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} [\cos(m-n)x - \cos(m+n)x]\,dx$$

$$= \frac{1}{2}\left[\frac{\sin(m-n)x}{m-n}\Bigg|_{-\pi}^{\pi} - \frac{\sin(m+n)x}{m+n}\Bigg|_{-\pi}^{\pi}\right] = 0.$$

Lastly, by the formula

$$\sin\alpha\cos\beta = \frac{\sin(\alpha-\beta) + \sin(\alpha+\beta)}{2}$$

for any integer m and n ($m \neq n$), we get

$$\int_{-\pi}^{\pi} \sin mx \sin nx\,dx = \frac{1}{2}\int_{-\pi}^{\pi} [\sin(m-n)x + \sin(m+n)x]\,dx$$

$$= \frac{1}{2}\left[-\frac{\cos(m-n)x}{m-n}\Bigg|_{-\pi}^{\pi} - \frac{\cos(m+n)x}{m+n}\Bigg|_{-\pi}^{\pi}\right] = 0.$$

For $m = n$, we will have

$$\int_{-\pi}^{\pi} \sin nx \cos nx\,dx = \int_{-\pi}^{\pi} \frac{\sin 2nx}{2}\,dx = -\frac{\cos 2nx}{2n}\Bigg|_{-\pi}^{\pi} = 0. \blacktriangleright$$

Remark. For $m = n$ the integrals of the products of the trigonometric functions under investigation will be

$$\int_{-\pi}^{\pi} \cos^2 nx\,dx = \int_{-\pi}^{\pi} \frac{1+\cos 2nx}{2}\,dx$$

$$= \frac{1}{2}\left(x\Big|_{-\pi}^{\pi} + \frac{\sin 2nx}{2n}\Bigg|_{-\pi}^{\pi}\right) = \pi,$$

$$\int_{-\pi}^{\pi} \sin^2 nx\,dx = \int_{-\pi}^{\pi} \frac{1-\cos 2nx}{2}\,dx = \pi.$$

16.2 Fourier Series for a Function with Period 2π

We would like to compute the coefficients a_0, a_n, b_n ($n = 1, 2, \ldots$) of the trigonometric series (16.1), knowing $f(x)$.

Theorem 16.2. *If*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{16.2}$$

holds for all x, and the series on the right-hand side converges uniformly on the interval $[-\pi, \pi]$, and hence, due to its periodicity, it converges on the entire real axis, then

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx \quad (n = 0, 1, 2, \ldots), \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx \quad (n = 1, 2, \ldots). \end{aligned} \tag{16.3}$$

◄ Since the terms of (16.2) are continuous functions on the interval $[-\pi, \pi]$ and (16.2) uniformly converges on that interval, we infer that $f(x)$ is integrable and so (16.3) makes sense. Moreover, (16.2) can be integrated term by term. We get

$$\int_{-\pi}^{\pi} f(x)\, dx = \frac{a_0}{2} \int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} \cos nx\, dx + b_n \int_{-\pi}^{\pi} \sin nx\, dx \right),$$

or

$$\int_{-\pi}^{\pi} f(x)\, dx = a_0 \pi.$$

From this follows the first of (16.3) for $n = 0$. We now multiply both sides of (16.2) by $\cos mx$, where m is an arbitrary natural number

$$\begin{aligned} f(x) \cos mx = \frac{a_0}{2} \cos mx + \sum_{n=1}^{\infty} (a_n \cos mx \cos nx \\ + b_n \cos mx \sin nx). \end{aligned} \tag{16.4}$$

Series (16.4) converges uniformly by Theorem 14.2, just like (16.2), therefore it can be integrated term by term, which gives

$$\int_{-\pi}^{\pi} f(x)\cos mx\,dx = \frac{a_0}{2}\int_{-\pi}^{\pi}\cos mx\,dx$$
$$+\sum_{n=1}^{\infty}\left(a_n\int_{-\pi}^{\pi}\cos mx\cos nx\,dx + b_n\int_{-\pi}^{\pi}\cos mx\sin nx\,dx\right).$$

The trigonometric system being orthogonal, all the integrals on the right-hand side are zero, save for one, which corresponds to $n = m$. Therefore,

$$\int_{-\pi}^{\pi} f(x)\cos mx\,dx = a_m\int_{-\pi}^{\pi}\cos^2 mx\,dx = a_m\pi,$$

hence

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos mx\,dx \quad (m = 1, 2, \ldots).$$

Likewise, multiplying both sides of (16.2) by $\sin mx$ and integrating from $-\pi$ to π, we obtain

$$\int_{-\pi}^{\pi} f(x)\sin mx\,dx = b_m\int_{-\pi}^{\pi}\sin^2 mx\,dx = b_m\pi,$$

hence

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin mx\,dx \quad (m = 1, 2, \ldots). \blacktriangleright$$

Now let $f(x)$ be an arbitrary function with period 2π and integrable on the interval $[-\pi, \pi]$. We do not know beforehand whether or not it can be represented as the sum of a certain convergent trigonometric series. But using (16.3) we can calculate a_n and b_n.

Definition. The trigonometric series

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx + b_n\sin nx),$$

whose coefficients a_0, a_n, and b_n are defined through $f(x)$ by the formulas

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx \quad (n = 0, 1, 2, \ldots),$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx \quad (n = 1, 2, \ldots),$$

is called the *Fourier series* of $f(x)$, and a_n, b_n, defined by these formulas are called the *Fourier coefficients* of $f(x)$.

Each function $f(x)$ integrable on $[-\pi, \pi]$ corresponds to its Fourier series

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{16.5}$$

i.e., a trigonometric series, whose coefficients are given by (16.3). If, however, we only require that $f(x)$ be integrable on $[-\pi, \pi]$, then, generally speaking, we cannot replace the symbol of correspondence in (16.5) by the symbol of equality.

Frequently, a function $f(x)$ needs to be expanded into a trigonometric series defined only on the interval $[-\pi, \pi]$ and consequently it is not periodic. A Fourier series can be written for such a function because the coefficients of the Fourier integral in (16.3) are calculated for $[-\pi, \pi]$. If, however, the function $f(x)$ is extended periodically along the x-axis, i.e., over the interval $(-\infty, +\infty)$, then we get a function $F(x)$ that has period 2π and coincides with $f(x)$ on the interval $[-\pi, \pi]$, i.e., $F(x) = f(x)$ for all x in $[-\pi, \pi]$. The function $F(x)$ is called the *periodic extension* of $f(x)$. $F(x)$ does not have a single value at $x \equiv \pm\pi, \pm 3\pi, \pm 5\pi, \ldots$ or at the points of discontinuity of $f(x)$ in $[-\pi, \pi]$.

The Fourier series for $F(x)$ will be identical to that for $f(x)$. If the Fourier series for $f(x)$ converges to $F(x)$, then the sum of the series, being a periodic function, will yield a periodic extension of $f(x)$ on $[-\pi, \pi]$ over the whole of the x-axis.

Thus when considering the Fourier series for $f(x)$ defined on $[-\pi, \pi]$, we are also considering the Fourier series for $F(x)$. It is sufficient, therefore, that the tests for convergence of a Fourier series be formulated only for periodic functions.

16.3 Sufficient Conditions for the Fourier Expansion of a Function

We will now find the sufficient test for convergence of a Fourier series.

Definition. A function $f(x)$ is called *piecewise monotone* on the interval $[a, b]$, if the interval can be broken up by a finite number of points $a < x_1 < x_2 < \ldots < x_{n-1} < b$ into intervals (a, x_1), (x_1, x_2), $\ldots$, (x_{n-1}, b), in each of which $f(x)$ is monotone, i.e., it is either nondecreasing or nonincreasing (see Fig. 16.1).

Examples. (1) The function $f(x) = x^2$ is piecewise monotone in the interval $(-\infty, +\infty)$ since the interval can be broken up into two intervals $(-\infty, 0)$ and $(0, +\infty)$, in the former it decreases and in the latter it increases.

(2) The function $f(x) = \cos x$ is piecewise monotone on the interval $[-\pi, \pi]$, since the interval can be divided into two intervals $(-\pi, 0)$ and $(0, \pi)$, in the first of which $\cos x$ increases from -1 to $+1$, in the second it decreases from $+1$ to -1.

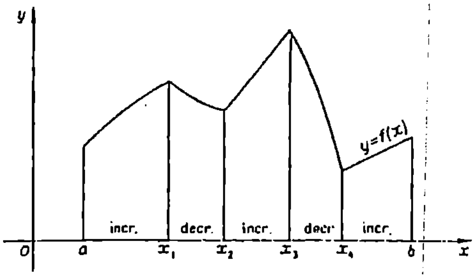


Fig. 16.1

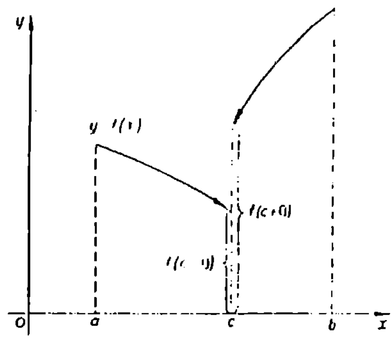


Fig. 16.2

If the function $f(x)$ is piecewise monotone and bounded on $[a, b]$ (i.e., $m \leqslant f(x) \leqslant M$), it can only have first-kind discontinuities in this interval. Let, for instance, $x = c$, $c \in [a, b]$ be a discontinuity of $f(x)$. Then there exist one-sided limits

$$\lim_{\substack{x \to c \\ x < c}} f(x) = f(c - 0) \quad \text{and} \quad \lim_{\substack{x \to c \\ x > c}} f(x) = f(c + 0).$$

This means that c is a first-kind discontinuity of $f(x)$ (Fig. 16.2).

Theorem 16.3. *If a periodic function $f(x)$ with period 2π is piecewise monotone and bounded on the interval $[-\pi, \pi]$, then its Fourier series converges at each point in the interval. The sum of the series*

$$S(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

obeys:

(1) $S(x) = f(x)$ *if* $-\pi < x < \pi$ *and* $f(x)$ *is continuous at* x;

(2) $S(x) = \frac{1}{2}[f(x + 0) + f(x - 0)]$ *if* $-\pi < x < \pi$, *and* x *is a discontinuity of* $f(x)$;

(3) $S(-\pi) = S(\pi) = \frac{1}{2}[f(-\pi + 0) + f(\pi - 0)]$.

Examples. (1) The function $f(x) = \pi - x$ with period 2π and defined in the interval $(-\pi, \pi)$, as is seen from its plot in Fig. 16.3, meets the conditions of the theorem. Therefore, it can be expanded into a Fourier series.

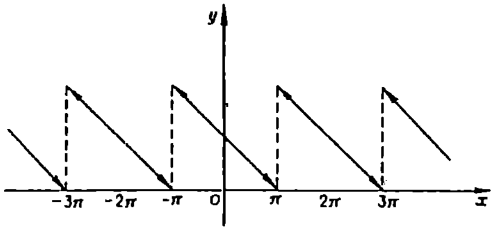


Fig. 16.3

◄ We find the Fourier coefficients for it, integrating by parts

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi}(\pi - x)\,dx = -\frac{(\pi - x)^2}{2\pi}\bigg|_{-\pi}^{\pi} = 2\pi,$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi}(\pi - x)\cos nx\,dx = \frac{1}{\pi}\int_{-\pi}^{\pi}(\pi - x)\,d\left(\frac{\sin nx}{n}\right)$$

$$= \frac{1}{\pi}(\pi - x)\frac{\sin nx}{n}\bigg|_{-\pi}^{\pi} + \frac{1}{\pi n}\int_{-\pi}^{\pi}\sin nx\,dx$$

$$= -\frac{\cos nx}{\pi n^2}\bigg|_{-\pi}^{\pi} = \frac{\cos(-n\pi) - \cos n\pi}{\pi n^2} = 0 \quad (n = 1, 2, \ldots),$$

$$b_n = -\frac{1}{\pi}\int_{-\pi}^{\pi}(\pi - x)\sin nx\,dx = \frac{1}{\pi}\int_{-\pi}^{\pi}(\pi - x)\,d\left(-\frac{\cos nx}{n}\right)$$

$$= -\frac{1}{\pi n}(\pi - x)\cos nx\bigg|_{-\pi}^{\pi} - \frac{1}{\pi n}\int_{-\pi}^{\pi}\cos nx\,dx$$

$$= \frac{1}{\pi n}2\pi\cos(-n\pi) - \frac{1}{\pi n^2}\sin nx\bigg|_{-\pi}^{\pi}$$

$$= \frac{2}{n}\cos n\pi = 2\frac{(-1)^n}{n} \quad (n = 1, 2, \ldots).$$

The Fourier series for the function has the form

$$\pi - x = \pi + 2\sum_{n=1}^{\infty}(-1)^n \frac{\sin nx}{n}, \quad -\pi < x < \pi. \blacktriangleright$$

(2) Expand the function $f(x) = \begin{cases} 0, & -\pi < x < 0, \\ x, & 0 \leqslant x < \pi, \end{cases}$ into a Fourier series (Fig. 16.4) on the interval $(-\pi, \pi)$.

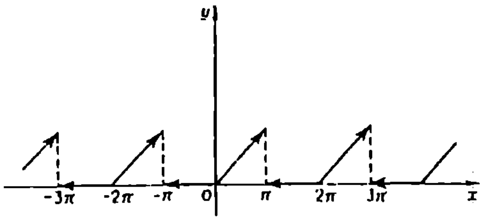


Fig. 16.4

◀ This function meets the conditions of the theorem.

Let us find the Fourier coefficients. By the additivity property of the definite integral,

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)dx = \frac{1}{\pi}\left(\int_{-\pi}^{0} f(x)dx + \int_{0}^{\pi} f(x)dx\right)$$

$$= \frac{1}{\pi}\left(\int_{-\pi}^{0} 0dx + \int_{0}^{\pi} xdx\right) = \frac{1}{\pi}\frac{x^2}{2}\bigg|_0^{\pi} = \frac{\pi}{2},$$

$$a_n = \frac{1}{\pi}\int_0^{\pi} x\cos nxdx = \frac{1}{\pi}\int_0^{\pi} x\, d\left(\frac{\sin nx}{n}\right)$$

$$= \frac{1}{\pi}\left(\frac{x\sin nx}{n}\bigg|_0^{\pi} - \frac{1}{n}\int_0^{\pi}\sin nxdx\right)$$

$$= \frac{1}{\pi n^2}\cos nx\bigg|_0^{\pi} = \frac{\cos n\pi - 1}{\pi n^2} = \frac{(-1)^n - 1}{\pi n^2}$$

$$= \begin{cases} -\dfrac{2}{\pi n^2} & \text{for} \quad n = 1, 3, 5, \ldots, \\ 0 & \text{for} \quad n = 2, 4, 6, \ldots, \end{cases}$$

$$b_n = \frac{1}{\pi}\int_0^\pi x \sin nx dx = \frac{1}{\pi}\int_0^\pi x\, d\left(-\frac{\cos nx}{n}\right)$$

$$= \frac{1}{\pi}\left(-\frac{x \cos nx}{n}\Bigg|_0^\pi + \frac{1}{n}\int_0^\pi \cos nx dx\right)$$

$$= -\frac{\cos n\pi}{n} = -\frac{(-1)^n}{n} = \frac{(-1)^{n+1}}{n} \quad (n = 1, 2, \ldots).$$

We thus arrive at

$$f(x) = \frac{\pi}{4} + \left(-\frac{2}{\pi}\frac{\cos x}{1^2} + \frac{\sin x}{1} - \frac{\sin 2x}{2} - \frac{2}{\pi}\frac{\cos 3x}{3^2}\right.$$

$$\left. + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} - \frac{2}{\pi}\frac{\cos 5x}{5^2} + \ldots\right)$$

$$= \frac{\pi}{4} + \sum_{n=1}^{\infty}\left[-\frac{2}{\pi}\frac{\cos(2n-1)x}{(2n-1)^2} + (-1)^{n+1}\frac{\sin nx}{n}\right]$$

for $-\pi < x < \pi$.

At the ends of the interval $[-\pi, \pi]$, i.e., at the points $x = -\pi$ and $x = \pi$, which are first-kind discontinuities of $f(x)$, we will have

$$f(-\pi) = f(\pi) = \frac{0 + \pi}{2} = \frac{\pi}{2}.$$

If we put $x = 0$, then

$$0 = \frac{\pi}{4} - \frac{2}{\pi}\left(\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \ldots\right),$$

hence

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots = \frac{\pi^2}{8}. \blacktriangleright$$

16.4. Fourier Expansions of Odd and Even Functions

Recall that a function $f(x)$ defined on the interval $[-l, l]$, where $l > 0$, is said to be even if $f(-x) = f(x)$ for all $x \in [-l, l]$.

The graph of an even function is symmetric about the ordinate axis.

A function $f(x)$ defined on the interval $[-l, l]$, where $l > 0$, is said to be odd if $f(-x) = -f(x)$ for all $x \in [-l, l]$.

The graph of an odd function is symmetric about the origin of coordinates.

Examples. (1) The function $f(x) = \cos x$ is even on the interval $[-\pi, \pi]$, since $\cos(-x) = \cos x$ for all x $[-\pi, \pi]$.

(2) The function $f(x) = \sin x$ is odd on the interval $[-\pi, \pi]$, since $\sin(-x) = -\sin x$ for all $x \in [-\pi, \pi]$.

(3) The function $x^2 - x$ is neither odd nor even on the interval $[-\pi, \pi]$, since

$$(-x)^2 - (-x) = x^2 + x \neq x^2 - x \quad \text{for} \quad x \in [-\pi, \pi], \quad x \neq 0.$$

Let a function $f(x)$ that meets the conditions of Theorem 16.3 be even on the interval $[-\pi, \pi]$, i.e., $f(-x) = f(x)$ for all $x \in [-\pi, \pi]$. We will then have $f(-x) \cos(-nx) = f(x) \cos nx$ for all $x \in [-\pi, \pi]$, i.e., $f(x) \cos nx$ is an even function, and $f(x) \sin nx$ is an odd function, since $f(-x) \sin(-nx) = -f(x) \sin nx$ for $x \in [-\pi, \pi]$. Therefore, the Fourier coefficients of the even function $f(x)$ will be

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx = \frac{2}{\pi} \int_{0}^{\pi} f(x) \cos nx\, dx \quad (n = 0, 1, 2, \ldots),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx = 0 \quad (n = 1, 2, \ldots).$$

Consequently, the Fourier series of an even function contains only cosines, i.e., has the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx.$$

If $f(x)$ is an odd function, i.e., $f(-x) = -f(x)$ for all $x \in [-\pi, \pi]$, then the product $f(x) \cos nx$ will be an odd function, and $f(x) \sin nx$ will be an even function. Therefore,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx = 0 \quad (n = 0, 1, 2, \ldots),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx = \frac{2}{\pi} \int_{0}^{\pi} f(x) \sin nx\, dx \quad (n = 1, 2, \ldots).$$

The Fourier series for an odd function thus contains only sines, i.e., it has the form

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx.$$

Examples. (1) Expand the function $f(x) = x^2$ into a Fourier series on the interval $-\pi \leqslant x \leqslant \pi$.

◄ Since the function is even and meets the conditions of Theorem 16.3, then its Fourier series has the form

$$x^2 = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx.$$

We find the Fourier coefficients

$$a_0 = \frac{2}{\pi}\int_0^{\pi} x^2 dx = \frac{2}{\pi}\frac{x^3}{3}\bigg|_0^{\pi} = \frac{2}{3}\pi^2.$$

If we integrate by parts twice, we will get

$$a_n = \frac{2}{\pi}\int_0^{\pi} x^2 \cos nx\, dx = \frac{2}{\pi}\int_0^{\pi} x^2\, d\left(\sin\frac{nx}{n}\right)$$

$$= \frac{2}{\pi}\left(x^2\frac{\sin nx}{n}\bigg|_0^{\pi} - \frac{2}{\pi}\int_0^{\pi} x \sin nx\, dx\right) = \frac{4}{n\pi}\int_0^{\pi} x\, d\left(\frac{\cos nx}{n}\right)$$

$$= \frac{4}{\pi n^2}\left(x\cos nx\bigg|_0^{\pi} - \int_0^{\pi}\cos nx\, dx\right) = \frac{4}{\pi n^2}\pi\cos n\pi = 4\frac{(-1)^n}{n^2},$$

$n = 1, 2, \ldots$.

And so the Fourier series for the function has the form

$$x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}(-1)^n\frac{\cos nx}{n^2}$$

or, if we unfold the formula,

$$x^2 = \frac{\pi^2}{3} - 4\left(\frac{\cos x}{1^2} - \frac{\cos 2x}{2^2} + \frac{\cos 3x}{3^2} + \ldots\right).$$

This relation is valid for all $x \in [-\pi, \pi]$, since at $x = \pm\pi$ the series has the sum that coincides with the values of $f(x) = x^2$, therefore

$$\frac{f(-\pi) + f(\pi)}{2} = \frac{\pi^2 + \pi^2}{2} = \pi^2 = f(\pi) = f(-\pi).$$

The graphs of $f(x) = x^2$ and of the sum of the resultant series are given in Fig. 16.5. ►

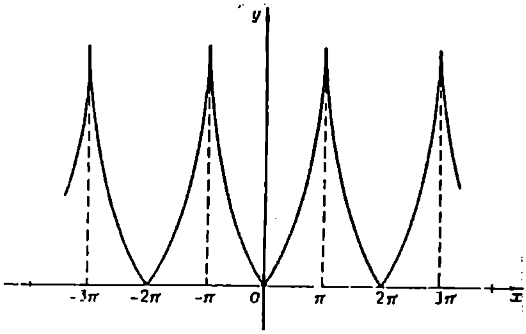


Fig. 16.5

Remark. Note that this Fourier series allows us to find the sums of some convergent number series. Namely at $x = 0$, we get

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \ldots = \frac{\pi^2}{12}.$$

For $x = \pi$ we get

$$\pi^2 = \frac{\pi^2}{2} - 4\left(-1 - \frac{1}{2^2} - \frac{1}{3^2} - \ldots\right),$$

whence

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots = \frac{\pi^2}{6}.$$

(2) Expand the function $f(x) = x$, $-\pi < x < \pi$, into a Fourier series on the interval $(-\pi, \pi)$.

◄ The function $f(x)$ meets the conditions of Theorem 16.3, and so it can be expanded into a Fourier series, which for this finite function has the form

$$x = \sum_{n=1}^{\infty} b_n \sin nx.$$

Integrating by parts gives the Fourier coefficients

$$b_n = \frac{2}{\pi}\int_0^{\pi} x \sin nx\, dx = \frac{2}{\pi}\int_0^{\pi} x\, d\left(-\frac{\cos nx}{n}\right)$$

$$= -\frac{2}{n\pi}\left(x\cos nx\Big|_0^\pi - \int_0^\pi \cos nx\,dx\right)$$

$$= -\frac{2}{n}\cos n\pi = -\frac{2}{n}(-1)^n = 2\frac{(-1)^{n+1}}{n} \quad (n = 1, 2, \ldots).$$

The Fourier series will thus be

$$x = 2\sum_{n=1}^{\infty}(-1)^{n+1}\frac{\sin nx}{n} = 2\left(\sin x - \frac{\sin 2x}{2} + \ldots\right).$$

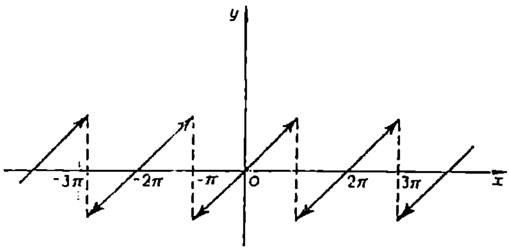


Fig. 16.6

This relation occurs for all $x \in (-\pi, \pi)$. At $x = \pm\pi$ the sum of the Fourier series, by Theorem 16.3, does not coincide with the values of $f(x) = x$, since the sum will be $[f(-\pi) + f(\pi)]/2 = (-\pi + \pi)/2 = 0$. Outside the interval $[-\pi, \pi]$ the sum is a periodic continuation of $f(x) = x$; its plot is given in Fig. 16.6. ►

16.5 Expansion of a Function Defined on the Given Interval into a Series of Sines and Cosines

Let a bounded piecewise monotone function $f(x)$ be defined on the interval $[0, \pi]$. If we in some way or another define it additionally on $[-\pi, 0]$, we can expand it into a Fourier series. For example, we can define $f(x)$ on $[-\pi, 0]$ so that $f(x) = f(-x)$. In that case, we say that $f(x)$ is "extended on the interval $[-\pi, 0]$ in an even manner". Its Fourier series will contain only cosines. If we define $f(x)$ on $[-\pi, 0]$ so that $f(x) = -f(-x)$, then we will obtain an odd function and we say that $f(x)$ is "extended on the interval $[-\pi, 0]$ in an odd manner". Its Fourier series will then contain only sines.

Each bounded piecewise monotone function $f(x)$ defined on $[0, \pi]$ can thus be expanded into a Fourier series in sines or cosines.

Examples. (1) Expand $f(x) = \pi - x$, $0 \leqslant x \leqslant \pi$, into a Fourier series in (a) sines and (b) cosines.

The function, when extended oddly or evenly onto $[-\pi, 0]$, will be bounded and piecewise monotone.

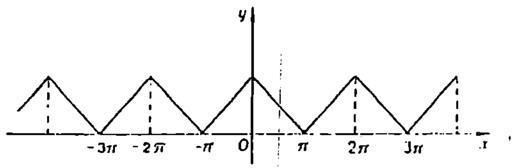


Fig. 16.7

◄ (a) Extend $f(x)$ onto the interval $[-\pi, 0]$ in an even manner (Fig. 16.7), then its Fourier series will be

$$\pi - x = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx,$$

where the Fourier coefficients are

$$a_0 = \frac{2}{\pi} \int_0^{\pi} (\pi - x)\,dx = -\frac{1}{\pi} (\pi - x)^2 \Big|_0^{\pi} = \pi,$$

$$a_n = \frac{2}{\pi} \int_0^{\pi} (\pi - x) \cos nx\,dx = \frac{2}{\pi} \int_0^{\pi} (\pi - x)\, d\left(\frac{\sin nx}{n}\right)$$

$$= \frac{2}{\pi} \left[(\pi - x) \frac{\sin nx}{n} \Big|_0^{\pi} + \frac{1}{n} \int_0^{\pi} \sin nx\,dx \right]$$

$$= -\frac{2}{\pi n^2} \cos nx \Big|_0^{\pi} = \frac{2}{\pi n^2} (1 - \cos n\pi)$$

$$= \frac{2}{\pi n^2} [1 - (-1)^n] = \begin{cases} \dfrac{4}{\pi n^2} & \text{for } n = 1, 3, 5, \ldots, \\ 0 & \text{for } n = 2, 4, 6, \ldots . \end{cases}$$

Thus,

$$\pi - x = \frac{\pi}{2} + \frac{4}{\pi} \left(\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \ldots \right),$$

where $0 \leqslant x \leqslant \pi$.

(b) Extend $f(x)$ on the interval $[-\pi, 0]$ in an odd manner (Fig. 16.8). Then its Fourier series will be

$$\pi - x = \sum_{n=1}^{\infty} b_n \sin nx,$$

where

$$b_n = \frac{2}{\pi}\int_0^{\pi} (\pi - x)\sin nx\, dx = \frac{2}{\pi}\int_0^{\pi}(\pi - x)\, d\left(-\frac{\cos nx}{n}\right)$$

$$= \frac{2}{\pi}\left[-(\pi - x)\frac{\cos nx}{n}\Big|_0^{\pi} - \frac{1}{n}\int_0^{\pi}\cos nx\, dx\right] = \frac{2}{\pi}\cdot\frac{\pi}{n} = \frac{2}{n},$$

for $n = 1, 2, \ldots$, and so

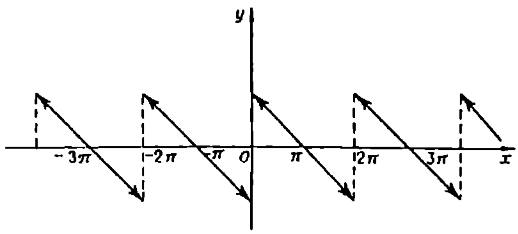


Fig. 16.8

$$\pi - x = 2\left(\frac{\sin x}{1} + \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \ldots\right)$$

for $0 < x \leqslant \pi$. ►

16.6 Fourier Series of a Function with Arbitrary Period

Let $f(x)$ be a periodic function with period $2l$, $l \neq 0$. To expand it into a Fourier series on $[-l, l]$, where $l > 0$, we change the variable $x = lt/\pi$. The function $F(t) = f(lt/\pi)$ will then be periodic in t with period 2π, i.e.,

$$F(t + 2\pi) = f\left[\frac{l}{\pi}(t + 2\pi)\right] = f\left(\frac{l}{\pi}t + 2l\right) = f\left(\frac{l}{\pi}t\right) = F(t)$$

and it can be expanded on the interval $[-\pi, \pi]$ into a Fourier series

$$F(t) = f\left(\frac{l}{\pi}t\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt),$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{lt}{\pi}\right) \cos nt\,dt \quad (n = 0, 1, 2, \ldots),$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\left(\frac{l}{\pi}t\right) \sin nt\,dt \quad (n = 1, 2, \ldots).$$

Return to the variable x, i.e., put $t = \pi x/l$, $dt = \frac{\pi}{l}\,dx$. We will obtain

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{l} + b_n \sin\frac{n\pi x}{l}\right),$$

$$a_n = \frac{1}{l}\int_{-l}^{l} f(x) \cos\frac{n\pi x}{l}\,dx \quad (n = 0, 1, 2, \ldots),$$

$$b_n = \frac{1}{l}\int_{-l}^{l} f(x) \sin\frac{n\pi x}{l}\,dx \quad (n = 1, 2, \ldots).$$

All the theorems for the Fourier series of functions with period 2π also hold for functions with any period $2l$. Specifically, we can here apply the sufficient test for the expandability of a function into a Fourier series.

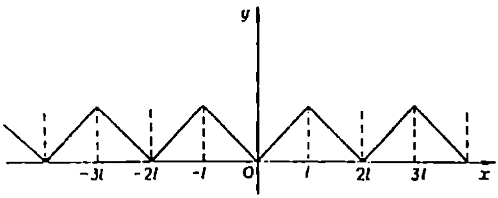


Fig. 16.9

Example. Expand into a Fourier series the function $f(x) = |x|$ with period $2l$ on the interval $[-l, l]$ (Fig. 16.9).

◄ Since the function is even, its Fourier series has the form

$$|x| = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{l},$$

where

$$a_0 = \frac{2}{l}\int_0^l x\,dx = l,$$

$$a_n = \frac{2}{l}\int_0^l x \cos\frac{n\pi x}{l}\,dx$$

$$= \frac{2}{l}\int_0^l x\,d\left(\frac{l}{n\pi}\sin\frac{n\pi x}{l}\right) = \frac{2}{\pi n}\int_0^l x\,d\left(\sin\frac{n\pi x}{l}\right)$$

$$= \frac{2}{n\pi}\left(x\sin\frac{n\pi x}{l}\Big|_0^l - \int_0^l \sin\frac{n\pi x}{l}\,dx\right)$$

$$= \frac{2}{n\pi}\frac{l}{\pi n}\cos\frac{n\pi x}{l}\Big|_0^l = \frac{2l}{n^2\pi^2}(\cos n\pi - 1)$$

$$= \frac{2l}{n^2\pi^2}[(-1)^n - 1] = \begin{cases} -\dfrac{4}{n^2\pi^2} & \text{for } n = 1, 3, 5, \ldots, \\ 0 & \text{for } n = 2, 4, 6, \ldots . \end{cases}$$

Substituting these values of the Fourier coefficients into the series gives

$$|x| = \frac{l}{2} - \frac{4l}{\pi^2}\left(\frac{\cos\frac{\pi x}{l}}{1^2} + \frac{\cos\frac{3\pi x}{l}}{3^2} + \frac{\cos\frac{5\pi x}{l}}{5^2} + \ldots\right)$$

for $-l \leqslant x \leqslant l$. ►

Note one important property of periodic functions.

If $f(x)$ has a period T and is integrable, then for any number a we will have

$$\int_a^{a+T} f(x)\,dx = \int_0^T f(x)\,dx,$$

i.e., the integral over an interval of length T has the same value regardless of the position of the interval on the number axis.

Indeed,

$$\int_a^{a+T} f(x)\,dx = \int_a^T f(x)\,dx + \int_T^{a+T} f(x)\,dx.$$

In the second integral we change the variable $x = t + T$, $dx = dt$. This gives

$$\int_T^{a+T} f(x)\,dx = \int_0^a f(t + T)\,dt = \int_0^a f(t)\,dt = \int_0^a f(x)\,dx,$$

hence

$$\int_a^{a+T} f(x)\,dx = \int_a^T f(x)\,dx + \int_0^a f(x)\,dx = \int_0^a f(x)\,dx + \int_a^T f(x)\,dx$$
$$= \int_0^T f(x)\,dx.$$

Geometrically, this property implies that if $f(x) \geqslant 0$ the areas hatched in Fig. 16.10 are equal.

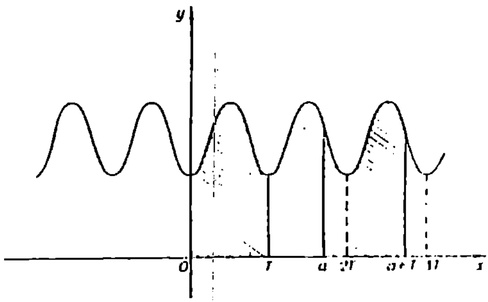


Fig. 16.10

Specifically, for $f(x)$ with period $T = 2\pi$ at $a = -\pi$ we will get

$$\int_{-\pi}^{\pi} f(x)\,dx = \int_0^{2\pi} f(x)\,dx.$$

Examples. (1) The function $f(x) = \sin^7 x$ is a periodic function with $T = 2\pi$. Therefore, without even taking the integrals we can state that for any a we will have

$$\int_a^{a+2\pi} \sin^7 x\,dx = \int_0^{2\pi} \sin^7 x\,dx = \int_{-\pi}^{\pi} \sin^7 x\,dx = 0,$$

since the function is odd. The property implies, in particular, that the Fourier coefficients of a periodic function $f(x)$ with period $2l$ can be worked out by

$$a_n = \frac{1}{l} \int_a^{a+2l} f(x) \cos \frac{n\pi x}{l} dx \quad (n = 0, 1, 2, \ldots), \tag{16.6}$$

$$b_n = \frac{1}{l} \int_a^{a+2l} f(x) \sin \frac{n\pi x}{l} dx \quad (n = 1, 2, \ldots), \tag{16.7}$$

where a is an arbitrary real number, since the functions $\cos \frac{n\pi x}{l}$ and $\sin \frac{n\pi x}{l}$ have period $2l$, and products of functions with period $2l$ will in turn be functions with period $2l$. ►

(2) Expand the function $f(x) = \begin{cases} \pi - x & \text{for } 0 < x < \pi, \\ 1 & \text{for } \pi < x < 2\pi, \end{cases}$ with period 2π (Fig. 16.11) into a Fourier series in the interval $0 < x < 2\pi$.

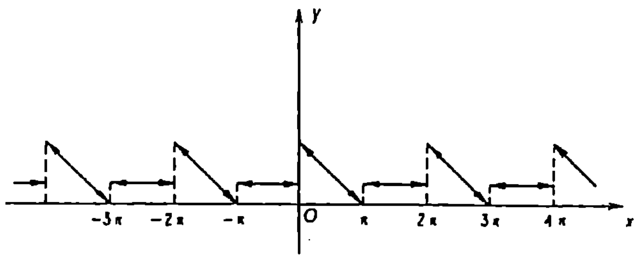


Fig. 16.11

◄ Find the Fourier coefficients. We put in (16.6) and (16.7) $a = 0$ and $l = \pi$. Then

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx = \frac{1}{\pi} \left[\int_0^{\pi} (\pi - x) dx + \int_{\pi}^{2\pi} dx \right] = \frac{\pi + 2}{2},$$

$$a_n = \frac{1}{\pi} \left[\int_0^{\pi} (\pi - x) \cos nx dx + \int_{\pi}^{2\pi} \cos nx dx \right]$$

$$= \frac{1}{\pi} \int_0^{\pi} (\pi - x) \cos nx dx = \begin{cases} \frac{2}{n^2 \pi} & \text{for } n = 1, 3, \ldots, \\ 0 & \text{for } n = 2, 4, \ldots, \end{cases}$$

$$b_n = \frac{1}{\pi}\left[\int_0^{\pi}(\pi - x)\sin nx\,dx + \int_{\pi}^{2\pi}\sin nx\,dx\right]$$

$$= \frac{1}{\pi}\left[\int_0^{\pi}(\pi - x)\,d\left(-\frac{\cos nx}{n}\right) - \frac{\cos nx}{n}\Big|_{\pi}^{2\pi}\right]$$

$$= \frac{1}{\pi}\left(-\frac{\pi - x}{n}\cos nx\Big|_0^{\pi} - \frac{1 - \cos n\pi}{n}\right)$$

$$= \frac{1}{\pi}\left[\frac{\pi}{n} + \frac{(-1)^n - 1}{n}\right] = \frac{1}{n\pi}\left[\pi + (-1)^n - 1\right]$$

$$= \begin{cases} \dfrac{\pi - 2}{n\pi} & \text{for } n = 1, 3, \ldots, \\[2ex] \dfrac{1}{n} & \text{for } n = 2, 4, \ldots . \end{cases}$$

The Fourier series will thus be

$$f(x) = \frac{\pi + 2}{4} + \frac{2}{\pi}\frac{\cos x}{1^2} + \frac{\pi - 2}{\pi}\frac{\sin x}{1} + \frac{\sin 2x}{2}$$

$$+ \frac{2}{\pi}\frac{\cos 3x}{3^2} + \frac{\pi - 2}{\pi}\frac{\sin 3x}{3} + \frac{\sin 4x}{4} + \ldots,$$

for $0 < x < 2\pi$.

At $x = \pi$ (first-kind discontinuity) we have

$$f(x) = \frac{f(\pi - 0) + f(\pi + 0)}{2} = \frac{0 + 1}{2} = \frac{1}{2}.$$

16.7 Complex Representation of Fourier Series

Suppose that $f(x)$ fulfils the sufficient conditions for Fourier expandability. It can then be represented on the interval $[-\pi, \pi]$ by a series of the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx). \tag{16.8}$$

Using the Euler formulas

$$e^{inx} = \cos nx + i \sin nx, \quad e^{-inx} = \cos nx - i \sin nx$$

we find that

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2}, \quad \sin nx = i\,\frac{e^{-inx} - e^{inx}}{2}.$$

Substituting these for cos nx and sin nx in (16.8) gives

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\,\frac{e^{inx} + e^{-inx}}{2} + ib_n\,\frac{e^{-inx} - e^{inx}}{2}\right)$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(\frac{a_n - ib_n}{2}\,e^{inx} + \frac{a_n + ib_n}{2}\,e^{-inx}\right). \tag{16.9}$$

We introduce the notation

$$\frac{a_0}{2} = c_0, \quad \frac{a_n - ib_n}{2} = c_n, \quad \frac{a_n + ib_n}{2} = c_{-n}.$$

Series (16.9) then becomes

$$f(x) = c_0 + \sum_{n=1}^{\infty}(c_n e^{inx} + c_{-n} e^{-inx}).$$

We now transform the right-hand side of this as follows:

$$f(x) = c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=1}^{\infty} c_{-n} e^{-inx} = c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=-1}^{-\infty} c_n e^{inx}$$

$$= \sum_{n=-\infty}^{-1} c_n e^{inx} + c_0 + \sum_{n=1}^{\infty} c_n e^{inx} = \sum_{n=-\infty}^{-1} c_n e^{inx} + \sum_{n=0}^{\infty} c_n e^{inx}.$$

Or in shorthand notation

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}. \tag{16.10}$$

The Fourier series (16.8) is thus represented in the complex form (16.10).

We will express c_n and c_{-n} in terms of integrals

$$c_n = \frac{a_n - ib_n}{2} = \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} f(x)\cos nx\,dx - i\int_{-\pi}^{\pi} f(x)\sin nx\,dx\right)$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(\cos nx - i\sin nx)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx.$$

Similarly,

$$c_{-n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{inx}dx.$$

These formulas for c_n and c_{-n} and c_0 and be combined into one expression

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx,$$

where $n = 0, \pm 1, \pm 2, \ldots$.

Coefficients c_n are called the complex Fourier coefficients of $f(x)$.

For a periodic function $f(x)$ with period $T = 2l$ $(l > 0)$ the Fourier series will be

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{l} + b_n \sin\frac{n\pi x}{l}\right)$$

and the complex form of the Fourier series becomes

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n e^{i\frac{n\pi x}{l}}, \tag{16.11}$$

where coefficients c_n are given by

$$c_n = \frac{1}{2l}\int_{-l}^{l} f(x)e^{-i\frac{n\pi x}{l}}dx \quad (n = 0, \pm 1, \pm 2, \ldots).$$

Series (16.10) and (16.11) are regarded as convergent for a given x if there exist the limits

$$\lim_{n\to\infty}\sum_{k=-n}^{n} c_k e^{ikx} \quad \text{and} \quad \lim_{n\to\infty}\sum_{k=-n}^{n} c_k e^{i\frac{k\pi x}{l}}.$$

Example. Expand into a complex Fourier series the periodic function

$$f(x) = \begin{cases} 0, & -\pi < x < 0 \\ 1, & 0 < x < \pi \end{cases} \quad \text{of period } 2\pi.$$

◄ The function meets the sufficient conditions for expandability into a Fourier series. Let

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n e^{inx}.$$

We find the complex Fourier coefficients:

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx = \frac{1}{2\pi}\int_{0}^{\pi} e^{-inx}dx$$

$$= \frac{1}{2\pi}\left(-\frac{1}{in}e^{-inx}\Big|_0^{\pi}\right) = \frac{1-e^{-in\pi}}{2\pi ni}$$

$$= \frac{1-\cos n\pi}{2\pi ni} = i\frac{(-1)^n - 1}{2\pi n} = \begin{cases} -\dfrac{i}{\pi n} & \text{for odd } n, \\ 0 & \text{for even } n, \end{cases}$$

or, in the shorthand form,

$$c_{2n-1} = -\frac{i}{\pi(2n-1)}.$$

Substituting c_n into the series gives

$$f(x) = -\frac{i}{\pi}\sum_{n=-\infty}^{+\infty}\frac{e^{i(2n-1)x}}{2n-1}, \quad -\pi < x < 0, \quad 0 < x < \pi.$$

Note that the series can be written as

$$f(x) = -\frac{i}{\pi}\sum_{n=-\infty}^{+\infty}\frac{e^{i(2n+1)x}}{2n+1}, \quad f(0) = \frac{1}{2}. \blacktriangleright$$

16.8 Fourier Series in General Orthogonal Systems of Functions

Orthogonal systems of functions. We denote by $L_2\,[a, b]$ the set of all (real) functions defined and integrable on the interval $[a, b]$, i.e., such that for them there exists the integral*)

$$\int_a^b f^2(x)dx < +\infty.$$

Specifically, functions $f(x)$ continuous on $[a, b]$ belong to $L_2\,[a, b]$.

Definition. A system of functions $\{\varphi_n(x)\}$, where $\varphi_n(n) \in L_2\,[a, b]$ is called *orthogonal* on $[a, b]$ if

$$(\varphi_m, \varphi_n) = \int_a^b \varphi_m(x)\varphi_n(x)dx = \begin{cases} 0 & \text{for } m \neq n, \\ \lambda_n > 0 & \text{for } m = n. \end{cases} \tag{16.12}$$

*) The Lebesgue integral.

Condition (16.12) implies, in particular, that neither of $\varphi_n(x)$ is identically zero.

We introduce the notation

$$\|\varphi_n\|^2 = (\varphi_n, \varphi_n) = \int_a^b \varphi_n^2(x)\,dx,$$

where $\|\varphi_n\|$ is said to be the *norm* of $\varphi_n(x)$.

If in the orthogonal system $\{\varphi_n(x)\}$ we have $\|\varphi_n\| = 1$ for all n, then the system of functions $\{\varphi_n(x)\}$ is said to be *orthonormal*.

If the system $\{\varphi_n(x)\}$ is orthogonal and $\varphi_n(x) \not\equiv 0$, then the system $\left\{\dfrac{\varphi_n(x)}{\|\varphi_n\|}\right\}$ is orthonormal.

Examples. (1) The trigonometric system

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \ldots,\ \cos nx,\ \sin nx,\ \ldots$$

is orthogonal on the interval $[-\pi, \pi]$.

The system of functions

$$\frac{1}{\sqrt{2\pi}},\ \frac{\cos x}{\sqrt{2\pi}},\ \frac{\sin x}{\sqrt{2\pi}},\ \ldots,\ \frac{\cos nx}{\sqrt{2\pi}},\ \frac{\sin nx}{\sqrt{2\pi}},\ \ldots$$

is an orthonormal system of functions on $[-\pi, \pi]$.

(2) The cosine system

$$1,\ \cos\frac{\pi x}{l},\ \cos\frac{2\pi x}{l},\ \ldots,\ \cos\frac{n\pi x}{l},\ \ldots$$

and the sine system

$$\sin\frac{\pi x}{l},\ \sin\frac{2\pi x}{l},\ \ldots,\ \sin\frac{n\pi x}{l},\ \ldots$$

are orthogonal on the interval $[0, l]$, but not orthonormal, since their norms

$$\left\|\cos\frac{n\pi x}{l}\right\| = \left\|\sin\frac{n\pi x}{l}\right\| = \sqrt{\frac{l}{2}} \neq 1.$$

(3) Polynomials given by

$$P_n(x) = \frac{1}{n!\,2^n}\frac{d^n(x^2-1)^n}{dx^n} \quad (n = 0, 1, 2, \ldots) \tag{16.13}$$

are called the *Legendre polynomials*.

At $n = 0$ we have $P_0(x) = 1$, at $n = 1$ we have $P_1(x) = \dfrac{1}{1!\,2}\dfrac{d}{dx}(x^2 - 1) = x$, at $n = 2$ we have $P_2(x) = \dfrac{3}{2}x^2 - \dfrac{1}{2}$, and so on.

It can be shown that the functions

$$\varphi_n(x) = \sqrt{\frac{2n+1}{2}}\, P_n(x) \quad (n = 0, 1, 2, \dots)$$

form an orthonormal system of functions on the interval $[-1, 1]$.

By way of example we show that Legendre polynomials are orthogonal.

◄ Let $m > n$. Integrating n times by parts we then find

$$\int_{-1}^{1} P_m(x)P_n(x)dx = \frac{1}{2^m 2^n m! n!} \int_{-1}^{1} \frac{d^n(x^2-1)^n}{dx^n} \frac{d^m(x^2-1)^m}{dx^m} dx$$

$$= \frac{1}{m!n!2^{m+n}} \left[\frac{d^{m-1}(x^2-1)^m}{dx^{m-1}} \cdot \frac{d^{n+1}(x^2-1)^n}{dx^{n+1}} \Bigg|_{-1}^{1} \right.$$

$$\left. - \int_{-1}^{1} \frac{d^{m-1}(x^2-1)^m}{dx^{m-1}} \frac{d^{n+1}(x^2-1)^n}{dx^{n+1}} dx \right]$$

$$= -\frac{1}{m!n!2^{m+n}} \int_{-1}^{1} \frac{d^{m-1}(x^2-1)^m}{dx^{m-1}} \frac{d^{n+1}(x^2-1)^n}{dx^{n+1}} dx$$

$$= \frac{(-1)^n}{m!n!2^{m+n}} \int_{-1}^{1} \frac{d^{m-n}(x^2-1)^m}{dx^{m-n}} \frac{d^{2n}(x^2-1)^n}{dx^{2n}} dx$$

$$= \frac{(-1)^n(2n)!}{m!n!2^{m+n}} \int_{-1}^{1} \frac{d^{m-n}(x^2-1)^m}{dx^{m-n}} dx = 0,$$

since all the derivatives up to the order of $m-1$ of the function $u_m = (x^2-1)^m$ vanish at the ends of the interval $[-1, 1]$. ►

Definition. A system of functions $\{\varphi_n(x)\}$ is called orthogonal in the interval (a, b) with weight $\varrho(x)$, if for all $n = 1, 2, \dots$ there exist integrals

(1) $$\int_a^b \varrho(x)\varphi_n^2(x)dx;$$

(2) $$\int_a^b \varrho(x)\varphi_m(x)\varphi_n(x)dx = \begin{cases} 0 & \text{for } m \neq n, \\ \lambda_n > 0 & \text{for } m = n. \end{cases}$$

It is assumed here that the weight function $\varrho(x) > 0$ everywhere in (a, b) except for, possibly, a finite number of points where $\varrho(x)$ is zero.

Examples. (1) A system of Bessel functions $\{J_\nu(\mu_i x)\}_{i=1}^{\infty}$ is orthogonal on the interval (0, 1) with weight $\varrho(x) = x$, i.e.,

$$\int_0^1 x J_\nu(\mu_j x) J_\nu(\mu_i x)\,dx = 0$$

for $i \neq j$. Here μ_i ($i = 1, 2, \ldots$) are zeros of the Bessel function $J_\nu(x)$.

(2) Consider the Chebyshev-Hermite polynomials, which can be defined as

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n} \quad (n = 0, 1, 2, \ldots). \tag{16.14}$$

Differentiating (16.14) gives

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2, \ldots.$$

It can be shown that Chebyshev-Hermite polynomials are orthogonal in the interval $(-\infty, +\infty)$ with weight $\varrho(x) = e^{-x^2}$, i.e.,

$$\int_{-\infty}^{+\infty} e^{-x^2} H_m(x) H_n(x)\,dx = 0 \quad \text{for} \quad m \neq n.$$

Fourier series for an orthogonal system. Let $\{\varphi_n(x)\}_{n=1}^{\infty}$ be an orthogonal system of functions in the interval (a, b) and let the series

$$c_1\varphi_1(x) + c_2\varphi_2(x) + \ldots + c_n\varphi_n(x) + \ldots \tag{16.15}$$

(c_i = const) be convergent in the interval to the function $f(x)$:

$$f(x) = c_1\varphi_1(x) + c_2\varphi_2(x) + \ldots + c_n\varphi_n(x) + \ldots.$$

Multiplying this by $\varphi_k(x)$ (here k is fixed) and integrating in x from a to b, we obtain by virtue of orthogonality of $\{\varphi_n(x)\}$

$$c_k = \frac{1}{\int_a^b \varphi_k^2(x)\,dx} \int_a^b f(x)\varphi_k(x)\,dx \tag{16.16}$$

or

$$c_k = \frac{(f, \varphi_k)}{\|\varphi_k\|^2} \quad (k = 1, 2, \ldots). \tag{16.16'}$$

This operation is, generally speaking, purely formal. Nevertheless, in certain cases, e.g., when series (16.15) converges uniformly, all the functions $\varphi_k(x)$ are continuous and the interval (a, b) is finite, this operation is quite legitimate. But at the moment it is the formal treatment that is important.

Thus, let the function $f(x) \in L_2\ [a, b]$ be defined. We find c_k by (16.16) and write

$$f(x) \sim c_1\varphi_1(x) + c_2\varphi_k(x) + \ldots + c_n\varphi_n(x) + \ldots . \tag{16.17}$$

The series on the right-hand side is called the *Fourier series* of $f(x)$ relative to the system $\{\varphi_n(x)\}$. Numbers c_n are called the *Fourier coefficients* of $f(x)$ for this system. The symbol $\sim$ in (16.17) only means that c_k are connected with $\varphi_k(x)$ by (16.16). Here it is not supposed that the series on the right converges at all, or less so to $f(x)$. It may well be asked: what are the properties of this series? In what sense does it represent the function $f(x)$?

Convergence in the mean. Definition. A sequence $\{f_n(x)\}, f_n(x) \in L_2\ [a, b]$, *converges* to the element $f \in L_2\ [a, b]$ *in the mean*, if

$$\lim_{n\to\infty} \int_a^b [f(x) - f_n(x)]^2 dx = 0$$

or, which is the same $\|f - f_n\| \to 0$, where $\|\ldots\|$ is the norm in the space $L_2\ [a, b]$.

Theorem 16.4. *If a sequence $\{f_n(x)\}$ converges uniformly, then it converges in the mean.*

◄ Suppose that $\{f_n(x)\}$ converges uniformly on $[a, b]$ to a function $f(x)$. This means that for any $\varepsilon > 0$ and all sufficiently large n we have

$$|f(x) - f_n(x)| < \varepsilon \quad \forall x \in [a, b],$$

hence

$$\int_a^b [f(x) - f_n(x)]^2 dx < \varepsilon^2(b - a).$$

And from this follows our statement. ►

The reverse is not true: the sequence $\{f_n(x)\}$ may converge in the mean to $f(x)$ but may not be uniformly convergent (here a and b are finite numbers). For example, let

$$f_n(x) = \frac{nx}{1 + n^2x^2} \quad (0 \leqslant x \leqslant 1).$$

It is easily seen that $\lim\limits_{n\to\infty} f_n(x) = 0\ \forall x \in [0, 1]$. But this convergence is nonuniform. There exists ε, e.g., $\varepsilon = 1/2$, such that at arbitrarily large n there is always a point in $[0, 1]$, namely $x = 1/n$, at which $f_n(x)$ is equal to $1/2$, i.e., $f_n(1/n) = 1/2$. And so by simply increasing n one cannot meet the condition $|f_n(x) - 0| < 1/2$ for all values of x, from 0 to 1, at the same time. In other words, already for $\varepsilon = 1/2$ there exists no N that would be suitable

for all $x \in [0, 1]$ at the same time (here the characteristic feature is the hump of height $1/2$ (Fig. 16.12) that shifts to the left with n).

On the other hand,

$$\int_0^1 [f_n(x) - 0]^2 dx = \int_0^1 \frac{n^2x^2}{(1 + n^2x^2)^2}\, dx$$

$$= -\frac{1}{2(1 + n^2)} + \frac{1}{2n} \tan^{-1} n \to 0, \quad n \to \infty$$

so that the sequence $\{f_n(x)\}$ converges in the mean to the function $f(x) \equiv 0$.

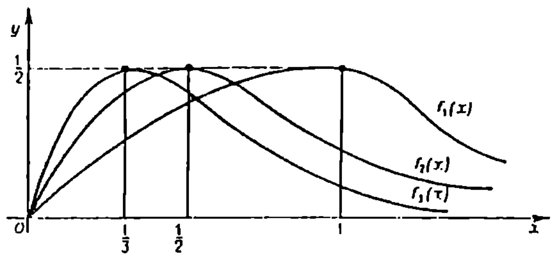


Fig. 16.12

Minimal property of Fourier coefficients. Bessel's inequality. Let $\{\varphi_n(x)\}$, $\varphi_n(x) \in L_2\,[a, b]$, be an orthonormal system of functions on $[a, b]$, i.e.,

$$\int_a^b \varphi_m(x)\varphi_n(x)\,dx = \begin{cases} 0, & m \neq n, \\ 1, & m = n, \end{cases}$$

and let $f(x) \in L_2\,[a, b]$.

Consider the linear combination

$$T_n(x) = \alpha_1\varphi_1(x) + \alpha_2\varphi_2(x) + \ldots + \alpha_n\varphi_n(x),$$

where $n \geqslant 1$ is a fixed integer. We find the values of the constants α_1, $\alpha_2, \ldots, \alpha_n$ for which the integral

$$\int_a^b [f(x) - T_n(x)]^2 dx \tag{*}$$

assumes a minimum value.

We denote by c_k the Fourier coefficients of $f(x)$ for the orthonormal system $\{\varphi_n(x)\}$

$$c_k = \int_a^b f(x)\varphi_k(x)\,dx \quad (k = 1, 2, \ldots).$$

By virtue of the system being orthonormal, we will have

$$\int_a^b [f(x) - T(x)]^2\,dx = \int_a^b [f^2(x) - 2f(x)T_n(x) + T_n^2(x)]\,dx$$

$$= \int_a^b f^2(x)\,dx - 2\sum_{k=1}^{n} c_k\alpha_k + \sum_{k=1}^{n} \alpha_k^2$$

$$= \int_a^b f^2(x)\,dx - \sum_{k=1}^{n} c_k^2 + \sum_{k=1}^{n} (\alpha_k - c_k)^2. \qquad (16.18)$$

The first two terms on the right-hand side of (16.18) are independent of α_k, and the third term is nonnegative. Therefore, the integral (*) assumes a minimum value at $\alpha_k = c_k$ ($k = 1, 2, \ldots$). The integral

$$\int_a^b [f(x) - T_n(x)]^2 dx$$

is termed the *mean square approximation of* $f(x)$ by the linear combination $T_n(x)$. The mean square approximation of $f(x) \in L\ [a, b]$ thus assumes a minimum value when $\alpha_k = c_k$ ($k = 1, 2, \ldots$), i.e., when $T_n(x)$ is the nth partial sum of the Fourier series of $f(x)$ for the system $\{\varphi_n(x)\}$: $T_n(x) = S_n(x)$. Setting $\alpha_k = c_k$, we obtain, by (16.6),

$$\int_a^b [f(x) - S_n(x)]^2 dx = \int_a^b f^2(x)dx - \sum_{k=1}^{n} c_k^2 \qquad (16.19)$$

or

$$\|f - S_n\| = \|f\|^2 - \sum_{n=1}^{\infty} c_k^2. \qquad (16.20)$$

Relation (16.20) is known as *Bessel's identity.* Its left-hand side is nonnegative, and so we obtain from it *Bessel's inequality*

$$\sum_{k=1}^{n} c_k^2 \leqslant \|f\|^2.$$

Since n is here arbitrary, we can represent Bessel's inequality in a stronger form

$$\sum_{k=1}^{\infty} c_k^2 \leqslant \|f\|^2, \qquad (16.21)$$

i.e., for any $f(x) \in L_2\,[a, b]$ the series of the squared Fourier coefficients of this function converges for the orthonormal system $\{\varphi_n(x)\}$.

Since the system

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos x}{\sqrt{2\pi}}, \quad \frac{\sin x}{\sqrt{2\pi}}, \quad \ldots,$$

is orthonormal on the interval $[-\pi, \pi]$, then in usual trigonometric notation (16.11) becomes

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty}(a_k^2 + b_k^2) \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)\,dx \qquad (16.22)$$

for any function $f(x)$ which is integrable when squared.

Corollary. If $f^2(x)$ is integrable, then $a_k \to 0$, $b_k \to 0$ as $k \to \infty$ (by the necessary condition for the convergence of the series on the left of (16.22)).

Parseval's formula. For certain systems $\{\varphi_n(x)\}$ the sign $\leqslant$ in (16.21) can be replaced (for all $f(x) \in L_2\,[a, b]$) by the equality sign. The resultant equality

$$\sum_{k=1}^{\infty} c_k^2 = \|f\|^2 \qquad (16.23)$$

is known as the *Parseval-Steklov formula* (*completeness condition*). It has the following meaning. The Bessel identity (16.20) allows the completeness condition to be written as

$$\lim_{n\to\infty} \|f - S_n\| = 0.$$

Condition (16.23) thus means that now the partial sums $S_n(x)$ of the Fourier series of $f(x)$ converge to $f(x)$ in the mean, i.e., by the norm of the space $L_2\,[a, b]$.

Definition. An orthonormal system $\{\varphi_n(x)\}$ is called a *complete system* in $L_2\,[a, b]$, if any function $f(x) \in L_2\,[a, b]$ can be approximated in the mean with any accuracy by the linear combination $\sum_{k=1}^{n} \alpha_k\varphi_k(x)$ with a sufficiently large number of terms, i.e., if for any $f(x) \in L_2\,[a, b]$ and any $\varepsilon > 0$ there is a natural number N_0 and numbers $\alpha_1, \alpha_2, \ldots, \alpha_{N_0}$, such that

$$\left\| f - \sum_{k=1}^{N_0} \alpha_k\varphi_k(x) \right\| < \varepsilon.$$

The foregoing suggests the following:

Theorem 16.5. *If an orthonormal system $\{\varphi_n(x)\}$ is complete in the space $L_2\,[a, b]$, then the Fourier series on any function $f(x) \in L_2\,[a, b]$ for that system converges to $f(x)$ in the mean, i.e., by the norm $L_2\,[a, b]$.*

Remark. It is possible to speak simply about complete orthogonal systems, because if system $\{\varphi_n(x)\}$ is orthogonal, the system $\left\{\dfrac{\varphi_n(x)}{\|\varphi_n\|}\right\}$ will be orthonormal.

It can be shown that the trigonometric series 1, $\cos x$, $\sin x$, ..., $\cos nx$, $\sin nx$, ... is complete in $L_2\,[-\pi, \pi]$. Whence the following theorem.

Theorem 16.6. *If $f(x) \in L_2\,[-\pi, \pi]$, then its trigonometric Fourier series converges in the mean.*

Closed systems. Completeness and closure. Definition. A system of functions $\{\varphi_n(x)\}$, $\varphi_n(x) \in L_2\,[a, b]$, is said to be *closed* if no nonzero function orthogonal to all the functions $\varphi_n(x)$ exists in $L_2\,[a, b]$.

There is no requirement in this definition for the system $\{\varphi_n(x)\}$ to be orthonormal.

Completeness and closure mean the same in $L_2\,[a, b]$ (the whole space) for orthonormal systems.

Exercises

Expand the following functions into the Fourier series in the interval $(-\pi, \pi)$

1. $f(x) = \begin{cases} 1 & \text{for} \quad -\pi < x < 0, \\ 0 & \text{for} \quad 0 < x < \pi. \end{cases}$ 2. $f(x) = \begin{cases} 2 & \text{for} \quad -\pi < x < 0, \\ 1 & \text{for} \quad 0 < x < \pi. \end{cases}$

3. $f(x) = \begin{cases} 0 & \text{for} \quad -\pi < x < 0, \\ 4x & \text{for} \quad 0 \leqslant x < \pi. \end{cases}$ 4. $f(x) = \begin{cases} -2x & \text{for} \quad -\pi < x < 0, \\ 3x & \text{for} \quad 0 \leqslant x < \pi. \end{cases}$

5. $f(x) = \pi + x$. 6. $f(x) = \begin{cases} 0 & \text{for} \quad -\pi < x < 0, \\ \pi - x & \text{for} \quad 0 < x < \pi. \end{cases}$

7. $f(x) = \sin^2 x$. 8. $f(x) = \dfrac{\pi^2 - 3x^2}{12}$.

9. $f(x) = |\sin x|$. 10. $f(x) = \dfrac{x}{2}$. 11. $f(x) = \sin \dfrac{x}{2}$.

12. Expand the function $f(x) = \pi - 2x$ given in the interval $(0, \pi)$ by extending the function into the interval $(-\pi, 0)$ (a) like an even function, (b) like an odd function.

13. Expand the function $f(x) = x^2$ given in the interval $(0, \pi)$ into the Fourier series in sines.

Expand into the Fourier series the functions

14. $f(x) = 3 - x$ in the interval $(-2, 2)$.
15. $f(x) = |x|$ in the interval $(-1, 1)$.
16. $f(x) = 2x$ in the interval $(0, 1)$.

Answers

1. $f(x) = \frac{1}{2} - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2n-1)x}{2n-1}$, $S(0) = \frac{1}{2}$, $S(\pm\pi) = \frac{1}{2}$.

2. $f(x) = \frac{3}{2} - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2n-1)x}{2n-1}$, $S(\pm\pi) = \frac{3}{2}$, $S(0) = \frac{3}{2}$.

3. $f(x) = \pi - \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2n-1)x}{(2n-1)^2} + 4\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sin nx}{n}$, $S(\pm\pi) = 2\pi$.

4. $f(x) = \frac{5}{4}\pi - \frac{10}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2n-1)x}{(2n-1)^2} + \sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sin nx}{n}$, $S(\pm\pi) = \frac{5}{2}\pi$.

5. $f(x) = \pi + 2\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sin nx}{n}$, $S(\pm\pi) = \pi$.

6. $f(x) = \frac{\pi}{4} + \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2n-1)x}{(2n-1)^2} + \sum_{n=1}^{\infty}\frac{\sin nx}{n}$, $S(\pm\pi) = \frac{\pi}{2}$.

7. $f(x) = \frac{1}{2} - \frac{1}{2}\cos 2x$. 8. $f(x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{\cos nx}{n^2}$.

9. $f(x) = \frac{2}{\pi} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos 2nx}{4n^2-1}$. 10. $f(x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sin nx}{n}$, $S(\pm\pi) = 0$.

11. $f(x) = \frac{8}{\pi}\sum_{n=1}^{\infty}(-1)^{n-1}\frac{n\sin nx}{4n^2-1}$.

12. (a) $f(x) = \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2n-1)x}{(2n-1)^2}$, (b) $f(x) = 2\sum_{n=1}^{\infty}\frac{\sin 2nx}{n}$.

13. $f(x) = 2\pi\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\sin nx}{n} - \frac{8}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2n-1)x}{(2n-1)^3}$.

14. $f(x) = 2 + \frac{4}{\pi}\sum_{n=1}^{\infty}(-1)^n\frac{\sin\frac{n\pi x}{2}}{n}$.

15. $f(x) = \frac{1}{2} - \frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{\cos(2n-1)\pi x}{(2n-1)^2}$. 16. $f(x) = 1 - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin 2n\pi x}{n}$.

Chapter 17

First-Order Ordinary Differential Equations

17.1 Basic Notions. Examples

An *ordinary differential equation* is an equation

$$F(x, y, y', \ldots, y^{(n)}) = 0 \tag{17.1}$$

relating an independent variable x, the desired function $y = y(x)$, and its derivatives $y'(x), y''(x), \ldots, y^{(n)}(x)$ (at least one derivative is needed). Here F is a specified function of its arguments.

In an ordinary differential equation the function $y = y(x)$ we seek is a function of one independent variable x; if it is a function of two (and more) independent variables, then the equation is called a partial differential equation. In what follows we will only deal with ordinary differential equations

The simplest differential equation is

$$y' = f(x), \tag{17.2}$$

where $f(x)$ is a known function that is continuous on a certain interval (a, b), and $y = y(x)$ is the desired function. We have already dealt with similar equations when we discussed integral calculus. There from a given function $f(x)$ we had to find its antiderivative (or primitive) $\mathscr{F}(x)$. Any function satisfying (17.2) is known to have the form

$$y = \mathscr{F}(x) + C,$$

where $\mathscr{F}(x)$ is some antiderivative for $f(x)$ on (a, b), and C is an arbitrary constant. The desired function $y = y(x)$ is thus not uniquely defined by (17.2).

The *order of a differential equation* is the order of the highest-order derivative present.

For example, $y' = xy^{100}$ is a differential equation of the first order; $y'' + \sin y = 0$ is a differential equation of the second order; $y^{(5)} + y'' + y = x + 1$ is a differential equation of the fifth order.

A *solution of a differential equation* of the nth order on the interval (a, b) is any function $y = \varphi(x)$, which has on this interval derivatives up to the nth order, such that the substitution of $y = \varphi(x)$ and its derivatives into the given equation turns the latter into an identity in x on (a, b).

For instance, the function $y = \sin x$ is a solution to the second-order differential equation $y'' + y = 0$ on the interval $(-\infty, +\infty)$, really, $y' = \cos x$, $y'' = -\sin x$. Substituting y and y'' into the equation gives $-\sin x + \sin x \equiv 0 \quad \forall x \in (-\infty, +\infty)$.

Problem. Find the coincident solutions of the differential equations (a) $y' = y^2 + 2x - x^4$ and (b) $y' = -y^2 - y + 2x + x^2 + x^4$.

The plot of a solution of a differential equation is called an *integral curve* of the equation.

Solving a differential equation is called the *integration of the differential equation.*

A host of problems in mathematics and other sciences, such as physics, chemistry, bioloby, etc., are modelled by differential equations.

Examples. (1) Find a curve such that the slope of the curve at each point would be equal to the ordinate of the point of tangency.

◄ Let $y = y(x)$ be the equation of the curve we seek. It is well known that $\tan \alpha = y'(x)$, and hence the properties of the curve are determined by the first-order differential equation $y'(x) = y(x)$. It is easily seen that $y = e^x$ is a solution of this equation. Also, it has the obvious solution $y \equiv 0$. Other solutions will be the functions $y = Ce^x$, where C is an arbitrary constant, so that the equation has an infinite number of solutions. ►

(2) Find the law of the rectilinear motion of a material point that travels with a constant acceleration a.

◄ We want to find the formula $s = s(t)$ that represents the path covered by the point as a function of time. As stated we have $d^2s/dt^2 = a$, which is a differential equation of the second order. We find

$$ds/dt = at + C_1, \quad s(t) = at^2/2 + C_1 t + C_2. \tag{*}$$

The constants can be determined by setting

$$s\Big|_{t=t_0} = s_0, \quad \frac{ds}{dt}\Bigg|_{t=t_0} = v_0.$$

Putting $t = t_0$ in the first of (*), we obtain $v_0 = at_0 + C_1$, hence $C_1 = v_0 - at_0$. From the second of (*) at $t = t_0$ we have

$$s_0 = \frac{at_0^2}{2} + C_1 t_0 + C_2 \quad \text{or} \quad s_0 = \frac{at_0^2}{2} + (v_0 - at_0)t_0 + C_2.$$

Hence

$$C_2 = s_0 - v_0 t_0 + \frac{at_0^2}{2}.$$

If then we substitute the values of C_1 and C_2 just found into the expression for $s(t)$, we will arrive at the well-known law of motion of a material point

with a constant acceleration:

$$s(t) = s_0 + v_0(t - t_0) + \frac{a(t - t_0)^2}{2}. \quad ▶$$

Let

$$F(x, y, y') = 0$$

be a differential equation of the first order. If we can solve it for y', we will obtain

$$y' = f(x, y), \tag{17.3}$$

where f is a given function of its arguments.

Along with (17.3) we can consider the equivalent differential equation

$$dy - f(x, y)\,dx = 0 \tag{17.3$'$}$$

or a more general form

$$M(x, y)\,dx + N(x, y)\,dy = 0, \tag{17.3$''$}$$

which can be obtained from (17.3′) multiplying it by a certain function $N(x, y) \not\equiv 0$ ($M(x, y)$ and $N(x, y)$ are known functions of their arguments).

Two differential equations $F_1(x, y, y') = 0$ and $F_2(x, y, y') = 0$ are called *equivalent* in a certain domain D of x, y, and y', if any solution $y(x)$ of one is also a solution of the other, and vice versa.

When handling differential equations one should see to it that the transformed equation is equivalent to the original one.

It follows from the above examples that a differential equation can have an infinite variety of solutions*). To isolate a definite solution of (17.3) we have to specify an *initial condition*, i.e., to state that at a certain value x_0 of x the desired function takes on a value y_0

$$y\big|_{x=x_0} = y_0 \quad \text{or} \quad y(x_0) = y_0. \tag{17.4}$$

Geometrically, this implies that we specify a point $M_0(x_0, y_0)$ through which the desired integral curve will have to pass.

The problem of finding a solution $y(x)$ of the equation (17.4) is called the *Cauchy problem* (initial-value problem) for equation (17.3).

---

*) The differential equation $(y' - 1)^2 + (x^2 - y^2)^2 = 0$ has only one solution $y = x$; the equation $(y')^2 + 1 = 0$ has no real solutions at all.

17.2 Solution of the Cauchy Problem for First-Order Differential Equations

Theorem 17.1 (existence and uniqueness of solution of the Cauchy problem). *Let*

$$y' = f(x, y) \tag{17.5}$$

be a differential equation and let $f(x, y)$ be defined in a certain domain D in the xy-plane. If there exists a neighbourhood Ω of a point $M_0(x_0, y_0) \in D$, where $f(x, y)$

(i) *is continuous in all arguments,*

(ii) *has a bounded partial derivative $\partial f/\partial y$,*

then there is an interval $(x_0 - h_0, x_0 + h_0)$ on the x-axis in which there exists a unique solution $y = \varphi(x)$ of equation (17.5), *such that at $x = x_0$ it is y_0* (Fig. 17.1).

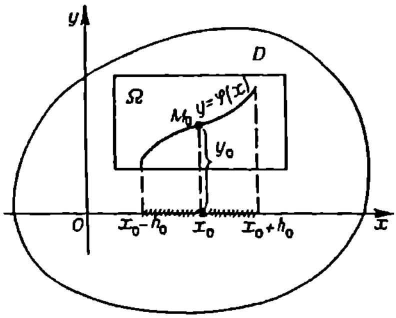


Fig. 17.1

Geometrically, this means that through the point $M_0(x_0, y_0)$ passes one and only one integral curve of (17.5).

Theorem 17.1 is local in nature: it only guarantees the existence of a unique solution $y = \varphi(x)$ of equation (17.5) within a sufficiently small neighbourhood of the point x_0. It follows from Theorem 17.1 that (17.5) has an infinite number of solutions (e. g., one solution whose plot passes through (x_0, y_0); another solution whose plot passes through (x_0, y_1), and so on).

Examples. (1) Let us take the equation $y' = x + y$.

◄ The function $f(x, y) = x + y$ is defined and continuous at all points in the xy-plane and everywhere $\partial f/\partial y = 1$. By Theorem 17.1, through each point (x_0, y_0) in the xy-plane passes one integral curve of that equation. ►

(2) Examine $y' = 3y^{2/3}$.

◄ The function $f(x, y) = 3y^{2/3}$ is defined and continuous in the entire xy-

plane; here $\partial f/\partial y = 2/y^{1/3}$ tends to infinity as y tends to zero and the second condition of Theorem 17.1 is violated on the x-axis. It can easily be verified that the function $y = (x + C)^3$, where C is any constant, is a solution of the given equation. Further, the equation has the obvious solution $y \equiv 0$. If we want to find a solution of the equation subject to the condition $y(0) = 0$, we will obtain an infinite number of such solutions, in particular the following (Fig. 17.2):

$$y \equiv 0, \quad y = \begin{cases} 0 & \text{for } x \leqslant 0, \\ x^3 & \text{for } x > 0, \end{cases} \qquad y = \begin{cases} x^3 & \text{for } x < 0, \\ 0 & \text{for } x \geqslant 0, \end{cases} \qquad y = x^3.$$

Thus, through each point on the x-axis pass at least two integral curves, and so uniqueness is violated on the axis.

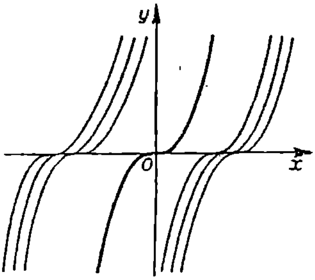


Fig. 17.2

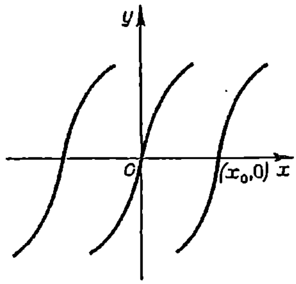


Fig. 17.3

If we take the point M_1 (1, 1) we will see that within its sufficiently small neighbourhood all the conditions of Theorem 17.1 are satisfied. Consequently, through this point within a small square Ω passes the only integral curve $y = x^3$ for $x > 0$ of the equation $y' = 3y^{2/3}$. If we take a sufficiently large square Ω (think how large), then we will no longer have a unique solution in it. This confirms the local character of Theorem 17.1

Theorem 17.1 provides *sufficient* conditions for the existence of a unique solution of the equation $y' = f(x, y)$. Namely: there may exist a unique solution $y = y(x)$ of the equation $y' = f(x, y)$ satisfying the condition $y|_{x = x_0} = y_0$, although at the point (x_0, y_0) one or both of the conditions (i) and (ii) of the theorem are not met.

For instance, for the equation $y' = 1/y^2$ we have $f(x, y) = 1/y^2$. On the x-axis the functions f and $\partial f/\partial y$ are discontinuous, and $\partial f/\partial y = -2/y^3$ tends to infinity as y tends to zero. But through each point $(x_0, 0)$ on the x-axis passes the only integral curve $y = \sqrt[3]{3(x - x_0)}$ (Fig. 17.3)

Remark. If we forego the boundedness of $\partial f/\partial y$, we will have the following theorem on the *existence* of the solution.

Theorem 17.2. *If the function $f(x, y)$ is continuous in a neighbourhood of the point (x_0, y_0), then the equation $y' = f(x, y)$ has at least one solution $y = \varphi(x)$, which at $x = x_0$ assumes the value y_0.*

Problems. (1) Find the integral curve for the equation $y' = \sin xy$ that passes through the point $O(0, 0)$.

(2) Find a solution for the Cauchy problem $dy/dx = \operatorname{sgn} y$, $y(x_0) = y_0$.

Definition. *A general solution* of the differential equation

$$y' = f(x, y) \tag{17.6}$$

in a certain domain Ω of the existence and uniqueness of the solution of the Cauchy problem is a uniparametric family S of functions $y = \varphi(x, C)$ that depend on x and one arbitrary constant C (parameter), such that

(1) at every permissible C the function $y = \varphi(x, C) \in S$ is a solution of (17.6), i.e.,

$$\varphi_x'(x, C) \equiv f(x, \varphi(x, C)), \quad x \in (x_0 - h, x_0 + h),$$

(2) whatever the initial condition $y\big|_{x = x_0} = y_0$, there exists a value of C_0 of C such that the solution $y = \varphi(x, C_0)$ will satisfy the initial condition

$$\varphi(x_0, C_0) = y_0.$$

It is also assumed that (x_0, y_0) belongs to the domain Ω of the existence and uniqueness of the solution of the Cauchy problem.

Example. Show that the equation $y' = 1$ has the general solution $y = x + C$, where C is an arbitrary constant.

◄ In this case, $f(x, y) \equiv 1$ and the conditions of Theorem 17.1 are satisfied everywhere. Accordingly, through each point (x_0, y_0) in the xy-plane passes one and only one integral curve of the equation.

We now test that $y = x + C$ meets the conditions (1) and (2) in the definition of the general solution. Really, for any C we have $y' = (x + C)' = 1$, so that $y = x + C$ is a solution of the given equation. If we require that at $x = x_0$ the solution would take on the value y_0, we arrive at $y_0 = x_0 + C$, whence $C_0 = y_0 - x_0$. The solution $y = x + y_0 - x_0$ or $y - y_0 = x - x_0$ complies with the initial condition. ►

A *particular solution* of the differential equation (17.6) is a solution derived from the general solution at some specific value of C (including $\pm\infty$). And so the *general solution* of this equation can be defined as a *set of all particular solutions.*

When integrating a differential equation we often arrive at the equation

$$\Phi(x, y, C) = 0, \tag{17.7}$$

which implicitly defines the general solution of the original equation.

Equation (17.7) is called the *general* (or *complete*) *integral* of (17.6).

Equation $\Phi(x, y, C_0) = 0$, where C_0 is some specific value of C, is called a *particular integral*.

In what follows we will sometimes say for short that a solution of an equation passes through a point $M_0(x_0, y_0)$ if M_0 lies on the curve of the solution.

Definition. A solution $y = \psi(x)$ of equation (17.6) is said to be *singular*, if the uniqueness property is violated at its every point, i.e., if through each point (x_0, y_0) of it in addition to this solution passes another solution of (17.6) that does not coincide with $y = \psi(x)$ within an arbitrarily small neighbourhood of (x_0, y_0).

The plot of a singular solution is called a *singular integral curve* of the equation. Geometrically, this is an *envelope* of the family of integral curves of the differential equation, defined by its general integral. (Recall that the *envelope* of a family of curves $\Phi(x, y, C) = 0$ is a curve that at every point is tangent to a certain curve of the family and each segment of which is tangent to an infinite set of curves of that family.)

If in a certain domain D on the xy-plane equation (17.6) satisfies the conditions of Theorem 17.1, then through each point $(x_0, y_0) \in D$ passes one and only one integral curve $y = \varphi(x)$ of the equation. This curve belongs to the uniparametric family $\Phi(x, y, C) = 0$ of curves which form the general integral of (17.6), and is obtained from that family at a specific value of C, i.e., it is a particular integral of (17.6). No other solutions passing through (x_0, y_0) are possible here.

For equation (17.6) to have a singular solution it is necessary that the conditions of Theorem 17.1 were not satisfied. Specifically, if the right-hand side of (17.6) is continuous in D, then a singular solution can only pass through points where the derivative $\partial f/\partial y$ becomes infinite. For example, for the equation

$$y' = 3y^{2/3} \tag{17.8}$$

the function $f = 3y^{2/3}$ is continuous everywhere, but the derivative $\partial f/\partial y$ tends to infinity at $y = 0$, i.e., on the x-axis in the xy-plane. Equation (17.8) has the general solution $y = (x + C)^3$, i.e., a family of cubic parabolas, and the obvious solution $y \equiv 0$ that passes through points where $\partial f/\partial y$ is unbounded. The solution $y \equiv 0$ is a singular one, since through each point of it pass both the cubic parabola and the straight line $y = 0$ (see Fig. 17.2). And so at each point of the solution $y \equiv 0$ the uniqueness property is violated. The singular solution $y \equiv 0$ does not follow from the solution $y = (x + C)^3$ at any numerical value of C (including $\pm\infty$).

From Theorem 17.1 we can only deduce necessary conditions for the singular solution. The set of the points where the derivative $\partial f/\partial y$ is unbounded, if this set is a curve, may happen to be not a singular solution,

if only for the fact that the curve, generally speaking, is not an integral curve of (17.6).

If, for example, instead of (17.8) we take the equation

$$y' = 3y^{2/3} + a, \quad a = \text{const}, \quad a \neq 0, \tag{17.9}$$

then on the line $y = 0$ the boundedness condition for $\partial f/\partial y$ is still violated, but this straight line clearly is not an integral curve of (17.9).

To sum up, to find singular solutions of (17.6) one must

(1) find the set of points where $\partial f/\partial y$ becomes infinite;

(2) if this set forms one or more curves, check whether or not they are integral curves of (17.6);

(3) if the curves are integral ones, check whether or not the uniqueness property is violated at all their points.

If these conditions are met, the curve in question is a singular solution of (17.6).

Problem. Find singular solutions of equation $y' = \sqrt{1 - y^2}$. Plot the curves.

17.3 Approximate Methods of Integration of the Equation $y' = f(x, y)$

Method of isoclines. Let

$$y' = f(x, y), \tag{17.10}$$

be a differential equation, where $f(x, y)$ in a certain domain D in the xy-plane meets the conditions of Theorem 17.1. This equation defines at each point (x, y) in D a value of y', i.e., the slope of the integral curve at that point. It is said that equation (17.10) defines in D a *direction field*. To construct it, we have at each point $(x_0, y_0) \in D$ to represent, using a certain segment, the direction of the tangent to the integral curve at that point, the direction being given by $f(x_0, y_0)$. The set of these segments gives a graphic picture of the direction field.

The problem of integration of the differential equation (17.10) can now be formulated as follows: find a curve such that its slope at each point coincided with the direction of the field at that point. This treatment of a differential equation and its integration gives us a graphic method of solving the equation.

We will construct the integral curves using isoclinic lines. An *isoclinic line*, or isocline, is a locus in the xy-plane where the slopes of the desired integral curves have the same direction ($y' = \text{const}$).

It follows from this definition that the family of isoclines of (17.10) is given by

$$f(x, y) = k,$$

where k is a parameter. If we give to k close numerical values, we can construct a sufficiently thick network of isoclines. Using these isoclines we can then approximately construct the integral curves of the equation.

Example. Integrate the equation $y' = x$ by the method of isoclines. ◄ The family of isoclines for the equation is defined by the equation $x = k$. Putting $k = 0, +1, -1, \ldots$, we obtain the isoclines $x = 0$, $x = 1$, $x = -1, \ldots$, from which we then construct the integral curves of the equation (Fig. 17.4). ►

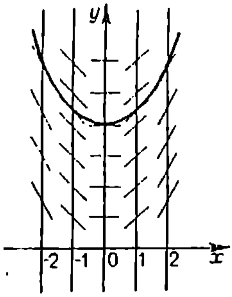


Fig. 17.4

The zero isocline $f(x, y) = 0$ defines the locus of extrema of integral curves (the line $x = 0$ in the example).

To improve the accuracy of construction it is recommended to find out whether the curve is convex up or convex down and the inflection points (if any) of the curves. To this end, y'' is determined by (17.10):

$$y'' = f_x' + f_y' y' = f_x'(x, y) + f_y'(x, y) f(x, y).$$

The sign of the right-hand part determines the sign of y'', i.e., determines whether the curve is convex up or convex down (see Vol. 1). The equation $f_x'(x, y) + f_y'(x, y) f(x, y) = 0$ defines the line which represents the locus of possible points of inflection of the integral curves.

In the above example $y'' = 1$ and therefore all the integral curves are convex downward and have no points of inflection.

Method of successive approximation. Let $y' = f(x, y)$ be a differential equation, where the function $f(x, y)$ satisfies the conditions of Theorem 17.1 in some domain D of x, y and let the point $(x_0, y_0) \in D$.

The solution of the Cauchy problem

$$dy/dx = f(x, y) \tag{17.11}$$

$$y(x_0) = y_0 \tag{17.12}$$

is equivalent to the solution of a certain integral equation, i.e., the equation with the unknown function under the integral sign. Indeed, let $y = y(x)$

be the solution of (17.11) specified in some neighbourhood $(x_0 - h_0, x_0 + h_0)$ of x_0 and satisfying the initial condition (17.12).

Then at $x \in (x_0 - h_0, x_0 + h_0)$ we have

$$\frac{dy}{dx} \equiv f(x, y(x)).$$

Integrating this identity in x and considering (17.12) we get

$$y(x) \equiv y_0 + \int_{x_0}^{x} f(t, y(t))\, dt, \quad x \in (x_0 - h_0, x_0 + h_0),$$

so that the solution $y(x)$ of the Cauchy problem satisfies the integral equation

$$y = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt. \tag{17.13}$$

Conversely, if a continuous function $y(x)$, $x \in (x_0 - h_0, x_0 + h_0)$ satisfies the integral equation (17.13) it is easy to verify that $y(x)$ is a solution of the Cauchy problem (17.11)-(17.12).

The solution $y = \varphi(x)$ of the integral equation (17.13) for all x sufficiently close to x_0 can be constructed through successive approximation by

$$\varphi(x) = \lim_{n \to \infty} \varphi_n(x),$$

where $\varphi_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, \varphi_n(t))\, dt$, $n = 0, 1, 2, \ldots$. Here $\varphi_0(t)$ can be any function continuous on the interval $[x_0 - h_0, x_0 + h_0]$, in particular $\varphi_0(t) = y_0 = \text{const}$.

Example. By successive approximation solve the Cauchy problem

$$\frac{dy}{dx} = y(x), \quad y(0) = 1.$$

We reduce the problem to the integral equation $y(x) = 1 + \int_0^x y(t)\, dt$.

If we take as our zeroth approximation the function $\varphi_0(x) = 1$, we successively find

$$\varphi_1(x) = 1 + \int_0^x \varphi_0(t)\, dt = 1 + x,$$

$$\varphi_2(x) = 1 + \int_0^x \varphi_1(t)\, dt = 1 + \int_0^x (1 + t)\, dt = 1 + x + \frac{x^2}{2},$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\varphi_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} \xrightarrow[n \to \infty]{} e^x.$$

It is easily seen that the solution is $y = e^x$. ►

Numerical methods. (1) *Euler method.* We want to find an approximate solution of the differential equation

$$\frac{dy}{dx} = f(x, y) \tag{17.14}$$

subject to the initial condition

$$y(x_0) = y_0. \tag{17.15}$$

Suppose that in a certain rectangle D: $\{|x - x_0| \leqslant A, |y - y_0| \leqslant B\}$ the function $f(x, y)$ is continuous and has continuous partial derivatives of sufficiently high orders in all the arguments, so that the solution of the Cauchy problem (17.14)-(17.15) exists, is unique and differentiable a sufficient number of times.

To solve (17.14)-(17.15) numerically means to construct a table of approximate values $y_1, y_2, ..., y_n$ of the solution at the points $x_1, x_2, ..., x_n$, respectively. The commonest points are $x_k = x_0 + kh$ ($k = 0, 1, ..., n$). Points x_k are called the mesh points, and $h > 0$ is called the step size of the mesh. Since by definition the derivative dy/dx is the limit of $\dfrac{y(x + h) - y(x)}{h}$ as $h \to 0$, then substituting this ratio for the derivative we, instead of (17.14), will obtain the difference equation (Euler difference scheme)

$$\frac{y_{k+1} - y_k}{h} = f(x_k, y_k) \quad (k = 0, 1, 2, \ldots) \tag{17.16}$$

or

$$y_{k+1} = y_k + hf(x_k, y_k) \quad (k = 0, 1, 2, \ldots). \tag{17.17}$$

By iteration we find $y_k = y(x_k)$, remembering that by (17.15) $y_0 = y(x_0)$ is a known quantity.

As a result, instead of the solution $y = y(x)$ we find the function $y_k = y(x_k)$ of the discrete argument x_k (mesh function), which yields an approximate solution of the problem (17.14)-(17.15). Geometrically, the desired integral curve $y = y(x)$ that passes through point $M_0(x_0, y_0)$ is replaced by an Euler broken line $M_0 M_1 M_2 \ldots$ with vertices at points $M_k(x_k, y_k)$ (Fig. 17.5).

The Euler method is a single-step method, which, to compute a point (x_{k+1}, y_{k+1}), requires a knowledge of the previous point (x_k, y_k) only. To estimate the error of the method in one step of the mesh we expand the exact solution $y = y(x)$ in a neighbourhood of the mesh points $x = x_k$ by the Taylor formula

$$\begin{aligned} y(x_{k+1}) &= y(x_k + h) = y(x_k) + y'(x_k)h + O(h^2) \\ &= y(x_k) + hf(x_k, y_k) + O(h^2). \end{aligned} \tag{17.18}$$

Comparison of (17.17) and (17.18) indicates that they coincide up to the terms of the first order in h, and the error of (17.17) is $O(h^2)$. We say then that the Euler method is of the first order.

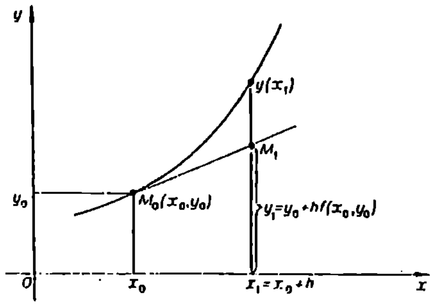


Fig. 17.5

Example. Using the Euler method, solve the Cauchy problem

$$\frac{dy}{dx} = y - x, \quad y(0) = 2$$

on the interval $[0, 0.5]$ with step $h = 0.1$.

◄ In this case, $f(x, y) = y - x$, $x_0 = 0$, $y_0 = 2$. Using (17.17), we obtain

$$y_{k+1} = y_k + hf(x_k, y_k)$$

we find consecutively

$$y_1 = y_0 + hf(x_0, y_0) = 2 + 0.1(2 - 0) = 2.2;$$

$$y_2 = y_1 + hf(x_1, y_1) = 2.2 + 0.1(2.2 - 0.1) = 2.41$$

and so on.

We tabulate the results:

k	x_k	y_k	$f(x_k, y_k)$	$hf(x_k, y_k)$	Exact solution $y = e^x + x + 1$
0	0	2.0000	2.0000	2.0000	2.0000
1	0.1	2.2000	2.1000	0.2100	2.2052
2	0.2	2.4100	2.2100	0.2210	2.4214
3	0.3	2.6310	2.3310	0.2331	2.6499
4	0.4	2.8641	2.4641	0.2464	2.8918
5	0.5	3.1105			3.1487

If we consider the Cauchy problem

$$\frac{dy}{dx} = y - x, \quad y(0) = 1$$

on any interval $[0, a]$ with any step $h > 0$, we will obtain $y_1 = 1 + h$, $y_2 = 1 + 2h$, $y_3 = 1 + 3h$, etc., so that the Euler broken line "straightens out" and coincides with the straight line $y = x + 1$, i.e., with the exact solution of the Cauchy problem.

(2) *Runge-Kutta method.* The Euler method is fairly simple but inaccurate. Accuracy can be improved by complicating the difference scheme, e.g., by the Runge-Kutta method.

We return to the Cauchy problem (17.14)-(17.15). We will again tabulate the approximate values $y_1, y_2, \ldots, y_n$ of the solution $y = y(x)$ of (17.14) at points $x_1, x_2, \ldots, x_n$ (mesh points).

Consider a scheme of equidistant mesh points $x_k = x_0 + kh$, where $h > 0$ is the step size of the mesh.

In the Runge-Kutta method the quantities y_{i+1} are computed by the following scheme:

$$y_{i+1} = y_i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

where

$$k_1 = f(x_i, y_i),$$
$$k_2 = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_1}{2}\right),$$
$$k_3 = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_2}{2}\right),$$
$$k_4 = f(x_i + h, y_i + hk_3).$$

17.4 Some Equations Integrable by Quadratures

A differential equation is said to be *integrable by quadratures* if its general solution (general integral) can be obtained as a result of a finite sequence of elementary operations with known functions and integrations of those functions. Such equations are relatively few in number. For example the equation $y' = x^2 + y^2$ is not integrable by quadratures. Consider some kinds of differential equations of the first order integrable by quadratures.

Separable equations. Equations of the type

$$f_1(y)\,dy = f_2(x)\,dx \tag{17.19}$$

are called *separated equations,* or *equations with separated variables.* Here $f_1(y), f_2(x)$ are known continuous functions of respective arguments.

Suppose that $y(x)$ is a solution of the equation. Then if we substitute $y(x)$ into (17.19), we will obtain an identity; and if we integrate it we will find the finite (not differential) equation

$$\int f_1(y)\,dy = \int f_2(x)\,dx + C, \tag{17.20}$$

which is satisfied by all the solutions of (17.19) (C is an arbitrary constant).

Conversely, each solution of (17.20) is a solution of the differential equation (17.19). Indeed, if some function $y(x)$, when substituted, turns (17.20) into an identity, then differentiating this identity shows that $y(x)$ also satisfies (17.19), i.e., is the general integral of this differential equation.

For instance, $x\,dx + y\,dy = 0$ is a separated equation. If we write it in the form $y\,dy = -x\,dx$ and integrate both parts, we will find the general integral of the equation: $x^2 + y^2 = C$.

An equation of the form

$$f_1(x)\varphi_1(y)\,dx = f_2(x)\varphi_2(y)\,dy, \tag{17.21}$$

where the coefficients at the differentials can be factored into components that depend only on x and only on y, is called a *separable differential equation,* since we can, by division by $\varphi_1(y)f_2(x) \neq 0$, reduce it to a separated equation:

$$\frac{f_1(x)}{f_2(x)}\,dx = \frac{\varphi_2(y)}{\varphi_1(y)}\,dy.$$

Example. Integrate the equation $(1 + y^2)x\,dx = (1 + x^2)y\,dy$.

◄ Dividing both sides of the equation by $(1 + y^2)(1 + x^2) \neq 0$ gives

$$\frac{x\,dx}{1 + x^2} = \frac{y\,dy}{1 + y^2}.$$

If then we integrate both sides of the resultant equality, we will get

$$\ln(1 + x^2) = \ln(1 + y^2) + \ln C, \quad \frac{1 + x^2}{1 + y^2} = C. \;►$$

Notice that division by $\varphi_1(y)f_2(x)$ may lead to a loss of solutions that turn $\varphi_1(y)f_2(x)$ into zero.

For example, separating the variables in $x\,dy = y\,dx$ gives $dy/y = dx/x$. Integration yields $\ln|y| = \ln|x| + \ln|C|$, whence $y = Cx$ (here C can assume both positive and negative values, but $C \neq 0$). Having divided by y we have lost the solution $y \equiv 0$, which can be included into the general solution $y = Cx$ if we allow C to take on the value $C = 0$.

If we assume that x and y may both enjoy equal rights, then we should

supplement the equation $dy/dx = y/x$, which makes no sense at $x = 0$, by the equation $dx/dy = x/y$, which has the obvious solution $x = 0$. In the general case, along with the differential equation

$$\frac{dy}{dx} = f(x, y), \tag{17.22}$$

we should also consider

$$\frac{dx}{dy} = f_1(x, y), \tag{17.22'}$$

where $f_1(x, y) = 1/f(x, y)$. In so doing, we should use (17.22′), where (17.22) makes no sense, and (17.22′) is meaningful.

By a change of variables we can reduce some differential equations to separable equations. Consider the equation of the form

$$\frac{dy}{dx} = f(ax + by + c), \tag{17.23}$$

where $f(z)$ is a continuous function, a, b, and c are constants. A substitution $z = ax + by + c$ yields the separable equation

$$\frac{dz}{dx} = a + b\frac{dy}{dx} = a + bf(z),$$

hence

$$\frac{dz}{a + bf(z)} = dx.$$

Integration gives $\int \frac{dz}{a + bf(z)} = x + C$. Changing z for $ax + by + c$, we find the general integral of (17.23).

Examples. (1) Integrate the equation $dy/dx = (x + y)^2$.

◄ We put $z = x + y$, then

$$\frac{dz}{dx} = 1 + \frac{dy}{dx} \quad \text{or} \quad \frac{dz}{dx} = 1 + z^2, \quad \text{hence} \quad \frac{dz}{1 + z^2} = dx.$$

Integrating gives $\tan^{-1} z = x + C$ or $z = \tan(x + C)$. Substituting $x + y$ for z, we obtain the general solution $y = \tan(x + C) - x$. ►

(2) It is common knowledge that the rate of radioactive decay is proportional to the amount x of the radioactive substance that has not yet decayed. Find the variation of x with time t, if at $t = t_0$ there was $x = x_0$ of substance.

◄ The process is described by the differential equation

$$\frac{dx}{dt} = -kx. \tag{*}$$

Here $k > 0$ is the decay constant, which is assumed to be known; the minus sign is to indicate that x decreases with t. Separating the variables in (*) and integrating yields

$$\ln|x| = -kt + \ln|C|, \quad x = Ce^{-kt}.$$

From the initial condition $x\big|_{t=t_0} = x_0$ we find $C = x_0 e^{kt_0}$, therefore

$$x(t) = x_0 e^{-k(t-t_0)}. \tag{**}$$

Any process (not only radioactive decay), in which the rate is proportional to the amount of substance that has not yet been involved, is described by equation (*).

The equation

$$\frac{dx}{dt} = kx, \quad k > 0, \tag{***}$$

which only differs by the sign on the right form (*), describes a multiplication process, e.g., the multiplication of neutrons in chain reactions or the multiplication of bacteria on the assumption that the rate of their multiplication is proportional to the available number of bacteria.

Equation (***) subject to the condition $x\big|_{t=t_0} = x_0$ has the solution $x(t) = x_0 e^{k(t-t_0)}$, which, unlike the solution of (**), grows with t.

The equations (*) and (***) can be merged to yield

$$\frac{dy}{dt} = ky, \quad k = \text{const}. \tag{****}$$

This equation is the simplest model of the dynamics of populations (the multitudes of individuals of one species of plant or animal organisms). Let $y(t)$ be the number of the members of the population at a time t. If we suppose that the rate of the variation of the population is proportional to the size of the population, then we arrive at equation (****). We then put $k = m - n$, where m is the coefficient of the relative birth rate, and n is the coefficient of the relative death rate, then $k > 0$ for $m > n$ and $k < 0$ for $m < n$.

If at $t = 0$ the size of the population is y_0, then equation (****) leads to the exponential law of population growth

$$y(t) = y_0 e^{kt}.$$

When $k < 0$, we have $y(t) \to 0$ as $t \to +\infty$, and when $k > 0$ we have $y(t) \to +\infty$ as $t \to +\infty$.

The assumption that m and n are constant does not hold for large populations. In fact, when a population is too large the resources available become depleted, which in turn reduces the birth rate and increases the death

rate. This circumstance can be described by simple laws

$$m = b_1 - b_2 y, \quad n = b_3 + b_4 y,$$

where b_i are positive constants ($i = 1, 2, 3, 4$). Then

$$k = m - n = b_1 - b_3 - (b_2 + b_4)y$$
$$= (b_2 + b_4)\left(\frac{b_1 - b_3}{b_2 - b_4} - y\right) = \alpha(A - y),$$

where $\alpha = b_2 + b_4$, $A = (b_1 - b_3)/(b_2 + b_4)$.

The population dynamics is then described by

$$\frac{dy}{dt} = \alpha(A - y)y.$$

This is the so-called *logistic equation,* a fundamental equation in demography and mathematical ecology. It is also used in the mathematical theory of propagation of rumours, deseases and in other problems of physiology and sociology. Separating the variables in the last equation, we obtain

$$\frac{dy}{(A - y)y} = \alpha\, dt, \quad y = \frac{A\,Ce^{A\alpha t}}{1 + Ce^{A\alpha t}}.$$

Letting $y(0) = y_0$, we will find the equation of the logistic curve

$$y(t) = \frac{A}{1 + \left(\dfrac{A}{y_0} - 1\right)e^{-A\alpha t}}.$$

When $\alpha > 0$ and $A > 0$, we find that $y(t) \to A$ as $t \to +\infty$. The logistic curve contains two parameters A and α; to find them we will have to know two additional values of $y(t)$ at some t_1 and t_2.

Equations homogeneous in x and y. We shall call a function $f(x, y)$ a *homogeneous function of the nth degree in x and y,* if for any admissible t we have

$$f(tx, ty) \equiv t^n f(x, y).$$

For instance, for the function $f(x, y) = x^2 - xy + y^2$ we have

$$f(tx, ty) = t^2x^2 - t^2xy + t^2y^2 = t^2(x^2 - xy + y^2) = t^2 f(x, y),$$

so that $f(x, y) = x^2 - xy + y^2$ is a homogeneous function in x and y of the second degree.

For the function $f(x, y) = y/x$ we have $f(tx, ty) = ty/tx = y/x = f(x, y)$, so that $f(x, y) = y/x$ is a homogeneous function of zeroth degree.

A differential equation of the first order $dy/dx = f(x, y)$ is said to be *homogeneous in z and y,* if the function $f(x, y)$ is a homogeneous function of zeroth degree in x and y.

Let $dy/dx = f(x, y)$ be a differential equation that is homogeneous in x and y. We put $t = 1/x$ in $f(tx, ty) = f(x, y)$ and obtain $f(x, y) = f(1, y/x)$, i.e., a homogeneous function of zeroth degree only depends on the ratio of arguments. If we denote $f(1, y/x)$ by $\varphi(y/x)$, we will see that a differential equation homogeneous in x and y can always be represented in the form

$$\frac{dy}{dx} = \varphi\left(\frac{y}{x}\right). \tag{17.24}$$

For an arbitrary continuous function φ the variables are inseparable. We introduce a new desired function $u(x) = y/x$, whence $y = xu$. Substituting $\frac{dy}{dx} = u + x\frac{du}{dx}$ into (17.24) gives

$$u + x\frac{du}{dx} = \varphi(u) \quad \text{or} \quad x\,du = [\varphi(u) - u]\,dx.$$

We now divide both sides of this by $x[\varphi(u) - u] \neq 0$ and integrate. The result will be

$$\int \frac{du}{\varphi(u) - u} = \ln|x| + \ln|C|.$$

Changing here u for its value y/x, we obtain the general integral of (17.24).

Examples. (1) Integrate the equation $\frac{dy}{dx} = \frac{y^2 + x^2}{xy}$.

◄ We have $dy/dx = y/x + x/y$. We put $y/x = u$, then $dy/dx = u + x(du/dx)$. The equation becomes

$$u + x\frac{du}{dx} = u + \frac{1}{u} \quad \text{or} \quad u\,du = \frac{dx}{x}.$$

Integrating gives $u^2 = \ln Cx^2$ or $y^2 = x^2 \ln Cx^2$. ►

(2) Find the shape of a mirror concentrating a parallel beam of light into a point.

◄ In the first place, the mirror should be a surface of revolution, since only for such a surface all the normals to the surface pass through the axis of revolution.

We will choose our coordinate system so that the rays were parallel to the x-axis and the reflected rays collected at the origin. We then find the shape of the cross-section of the mirror by the xy-plane. Let the cross-section equation be $y = \varphi(x)$ (Fig. 17.6).

We will draw a tangent line BN to the cross-section at a point $M(x, y)$, where the ray L touches the surface, its angle formed by BN and the x-axis will be α. Let now N be a point where the tangent line cuts the x-axis. By the reflection law, the angles NMO and BML must be equal. It is easily

seen that the angle MOP is 2α. Since $\tan\alpha = y'$, $\tan 2\alpha = y/x$ and $\tan 2a = \dfrac{2\tan\alpha}{1 - \tan^2\alpha}$, then at any point on the curve $y = \varphi(x)$ holds the relation $y/x = \dfrac{2y'}{1 - (y')^2}$, i.e., the differential equation that defines the required path of the ray.

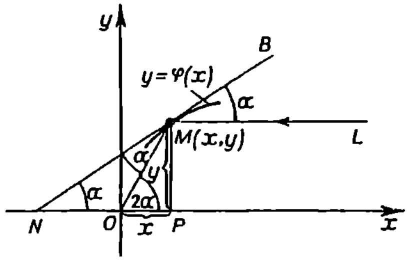


Fig. 17.6

Solving the equation for the derivative, we will obtain two homogeneous equations:

$$y' = \frac{-x + \sqrt{x^2 + y^2}}{y}, \quad y' = \frac{-x - \sqrt{x^2 + y^2}}{y}.$$

The first of these is transformed by a change $y/x = u$ to

$$u + xu' = \frac{-1 + \sqrt{1 + u^2}}{u} \quad \text{or} \quad \frac{u\,du}{1 + u^2 - \sqrt{1 + u^2}} = -\frac{dx}{x}.$$

Integrating, we will find

$$-\ln|x| + \ln C = \int \frac{u\,du}{1 + u^2 - \sqrt{1 + u^2}} = \int \frac{d\left(\sqrt{1 + u^2} - 1\right)}{\sqrt{1 + u^2} - 1}$$
$$= \ln\left(\sqrt{1 + u^2} - 1\right).$$

Taking antilogs and substituting y/x for u, we obtain after some algebra

$$y^2 = 2Cx + C^2 \quad \text{or} \quad y^2 = 2C(x + C/2).$$

The resultant equation defines in the xy-plane a family of parabolas symmetric about the x-axis; the foci of these parabolas coincide with the origin. We fix C and rotate a parabola about the x-axis to obtain the paraboloid of revolution $y^2 + z^2 = 2C(x + C/2)$. And so a mirror in the

shape of a paraboloid of revolution solves the problem. This property is used in flood lights. ►

Remark. If $\varphi(u) - u \equiv 0$, then equation (17.24) has the form $dy/dx = y/x$ and is integrated by separation of variables. Its general solution is $y = Cx$. If $\varphi(u) - u$ goes to zero at $u = u_0 = \text{const}$, then there also exists the solution $u = u_0$ or $y = u_0 x$ (a straight line passing through the origin). That there exists the solution $u = u_0$ is immediately seen in that case if we write the differential equation in the form $x\,du = [\varphi(u) - u]\,dx$.

Consider some equations that can be reduced to homogeneous ones. The equation

$$\frac{dy}{dx} = \frac{ax + by + c}{a_1 x + b_1 y + c_1}, \tag{17.25}$$

where a, b, c, a_1, b_1, and c_1 are constants, at $c = c_1 = 0$ is homogeneous. Suppose now that at least one of c and c_1 is nonzero. We should here distinguish two cases.

(1) The determinant $\begin{vmatrix} a & b \\ a_1 & b_1 \end{vmatrix}$ is nonzero. Introduce new variables ξ and η by

$$x = \xi + h, \quad y = \eta + k,$$

where h and k are as yet indefinite constants; then $dx = d\xi$, $dy = d\eta$. Equation (17.25) then becomes

$$\frac{d\eta}{d\xi} = \frac{a\xi + b\eta + ah + bk + c}{a_1\xi + b_1\eta + a_1 h + b_1 k + c_1}.$$

If we choose h and k as solutions of the system of linear algebraic equations

$$\begin{cases} ah + bk + c = 0, \\ a_1 h + b_1 k + c_1 = 0, \end{cases} \tag{17.26}$$

we obtain the homogeneous (in ξ and η) equation

$$\frac{d\eta}{d\xi} = \frac{a\xi + b\eta}{a_1\xi + b_1\eta}.$$

Replacing in its general integral ξ by $x - h$ and η by $y - k$, we will find the general inegral of (17.25).

(2) The determinant $\begin{vmatrix} a & b \\ a_1 & b_1 \end{vmatrix}$ is zero. System (17.26) in the general case has no solutions and the method just described is unsuitable. But in

this case $a_1/a = b_1/b = \lambda$, i.e., equation (17.25) has the form

$$\frac{dy}{dx} = \frac{ax + by + c}{\lambda(ax + by) + c_1}$$

and can be reduced to a separable equation by the substitution $z = ax + by$.

In a similar manner, we integrate the equation

$$\frac{dy}{dx} = f\left(\frac{ax + by + c}{a_1x + b_1y + c_1}\right),$$

where $f(w)$ is a continuous function.

Linear differential equations. A *linear differential equation of the first order* is an equation linear in an unknown function and its derivative. In the general case, it has the form

$$A(x)\frac{dy}{dx} + B(x)y = f(x), \tag{17.27}$$

where the coefficients $A(x)$ and $B(x)$ and the right-hand side $f(x)$ are taken to be defined on a certain interval (α, β).

If $f(x) \equiv 0$ on (α, β), this equation is called *homogeneous*, otherwise it is called *inhomogeneous*. Assuming $A(x) \neq 0$ on (α, β) and dividing both sides of (17.27) by $A(x)$, we reduce (17.27) to the form

$$\frac{dy}{dx} + p(x)y = q(x), \tag{17.28}$$

where $p(x) = B(x)/A(x)$, $q(x) = f(x)/A(x)$.

Theorem 17.3. *If functions $p(x)$ and $q(x)$ are continuous on $[a, b] \in (\alpha, \beta)$, then equation (17.28) always has a unique solution that satisfies the initial condition $y|_{x=x_0} = y_0$, where the point (x_0, y_0) belongs to the band $a < x < b$, $-\infty < y < +\infty$.*

◄ We solve (17.28) for y'

$$y' = -p(x)y + q(x).$$

Here the right-hand side meets all the conditions of Theorem 17.1: it is continuous in the variables x and y and has the bounded partial derivative $\partial f/\partial y = -p(x)$ in the band. This proves the statement. ►

A linear homogeneous equation corresponding to (17.28) has the form

$$\frac{dy}{dx} + p(x)y = 0. \tag{17.29}$$

It is integrated by separating the variables: $dy/y = -p(x)\,dx$. Hence

$$\ln|y| = -\int p(x)\,dx + \ln|C| \quad \text{or} \quad y = Ce^{-\int p(x)\,dx}. \tag{17.30}$$

Dividing by y we lose the solution $y \equiv 0$. However, it can be included into the family of solutions (17.30), if we assume that C can be zero. Formula (17.30) gives the general solution of (17.29) in the band $a < x < b$, $-\infty < y < +\infty$.

The inhomogeneous linear equation (17.28) can be integrated using the so-called *method of variation of constants*. It consists in the following. We at first integrate the homogeneous equation

$$\frac{dy}{dx} + p(x)\,y = 0,$$

whose general solution has the form $y = Ce^{-\int p(x)\,dx}$, where C is an arbitrary constant.

We seek a solution of (17.28) in the form

$$y = C(x)\,e^{-\int p(x)\,dx}, \tag{17.31}$$

where $C(x)$ is a new unknown function.

Calculating the derivative dy/dx and substituting the values of dy/dx and y into the original equation (17.28) gives

$$\frac{dC}{dx} = q(x)\,e^{\int p(x)\,dx}, \quad C(x) = \int q(x)\,e^{\int p(x)\,dx}dx + C,$$

where C is a constant of integration. Hence

$$y = C(x)\,e^{-\int p(x)\,dx} = Ce^{-\int p(x)\,dx} + e^{-\int p(x)\,dx}\int q(x)\,e^{\int p(x)\,dx}dx. \tag{17.32}$$

This is the general solution of the linear inhomogeneous differential equation (17.28).

It is seen from (17.32) that the general solution of (17.28) is the sum of the general solution of the corresponding homogeneous equation and the particular solution of (17.28) that follows from (17.32) at $C = 0$, i.e.,

$$y_{\text{g.i.}} = y_{\text{g.h.}} + y_{\text{p.i}}.$$

In (17.32) the indefinite integrals can be replaced by definite integrals with a variable upper limit

$$y(x) = e^{-\int_{x_0}^{x} p(x)\,dx}\left[C + \int_{x_0}^{x} q(x)\,e^{\int_{x_0}^{x} p(x)\,dx}dx\right].$$

Here $C = y(x_0) = y_0$, therefore the general solution of (17.28) can be written as

$$y = e^{-\int_{x_0}^{x} p(x)\,dx}\left[y_0 + \int_{x_0}^{x} q(x)\,e^{\int_{x_0}^{x} p(x)\,dx}dx\right], \tag{17.33}$$

where the role of an arbitrary constant is played by the initial value y_0 of the desired function $y(x)$.

Formula (17.33) is the general solution of (17.28) in the *Cauchy form*. It follows that if $p(x)$ and $q(x)$ are defined and continuous on the interval $-\infty < x < +\infty$, then the solution $y(x)$ of (17.28) with any initial data $y(x_0) = y_0$ will be continuous and even continuously differentiable at all finite x, so that the integral curve passing through any point (x_0, y_0) will be a smooth curve on the interval $-\infty < x < +\infty$.

Examples. (1) Integrate the equation

$$\frac{dy}{dx} + y\cos x = 2\cos x. \qquad (*)$$

◄ We will integrate the homogeneous equation $dy/dx + y\cos x = 0$ that corresponds to the original one by separating the variables

$$y = Ce^{-\sin x}.$$

We will seek the solution of the original equation in the form

$$y = C(x)e^{-\sin x}, \qquad (**)$$

where $C(x)$ is unknown function. We find dy/dx and substitute it and y into $(*)$

$$\frac{dy}{dx} = \frac{dC}{dx}e^{-\sin x} - C(x)e^{-\sin x}\cos x,$$

$$\frac{dC}{dx}e^{-\sin x} - C(x)e^{-\sin x}\cos x + C(x)e^{-\sin x}\cos x = 2\cos x,$$

$$\frac{dC}{dx} = 2\cos x\, e^{\sin x}; \quad C(x) = 2e^{\sin x} + C,$$

where C is the constant of integration.

From $(**)$ we find the general solution of $(*)$

$$y(x) = Ce^{-\sin x} + 2.$$

It is easy to see the particular solution of the inhomogeneous equation $(*)$. In general, if the particular solution of the linear inhomogeneous equation can be "figured out", then this substantially simplifies the search for its general solution. ►

(2) Consider the phenomenon that occurs on closing a d.c. circuit.

◄ If R is the resistance of the circuit, E the external e.m.f., then the current $I = I(t)$ will grow gradually from zero to a finite stationary value E/R.

Let L be the self-inductance of the circuit, such that whenever the current in the circuit changes, an e.m.f. emerges in the circuit such that its magnitude is $L\dfrac{dI}{dt}$ and its direction is opposite to that of the external

e.m.f. By Ohm's law, at each moment of time t the product of the current by the resistance equals the real e.m.f., and we get

$$IR = E - L\frac{dI}{dt} \quad \text{or} \quad \frac{dI}{dt} + \frac{R}{L}I = \frac{E}{L}. \tag{$*$}$$

$(E,\ L,\ R = \text{const.})$

Equation $(*)$ is a linear inhomogeneous equation in $I(t)$. It is easily seen that its particular solution is the function $I_p(t) = E(R)$. The general solution of the corresponding homogeneous equation is $I_g(t) = Ce^{-(R/L)t}$. And so the general solution of the inhomogeneous equation $(*)$ is

$$I(t) = Ce^{-\frac{R}{L}t} + \frac{E}{R}.$$

At $t = 0$ we have $I(0) = 0$, therefore $C = -E/R$, so that we arrive at

$$I(t) = \frac{E}{R}\left(1 - e^{-\frac{R}{L}t}\right).$$

It is seen that on closing the circuit as $t \to +\infty$ the current asymptotically tends to its stationary value E/R. ►

The linear inhomogeneous differential equation (17.28) can also be integrated using the following trick.

We will seek the solution $y(x)$ of (17.28) in the form

$$y(x) = u(x)\,v(x), \tag{17.34}$$

where $u(x)$ and $v(x)$ are unknown functions, one of which, say $v(x)$, can be arbitrary. Substituting $y(x)$ in the form (17.34) into (17.28) we obtain after some algebra

$$u'v + (v' + p(x)v)u = q(x). \tag{17.35}$$

We choose as $v(x)$ any particular solution $v(x) \not\equiv 0$ of the equation $v' + p(x)v = 0$. Then, by (17.35), we will obtain for $u(x)$ the equation

$$v\frac{du}{dx} = q(x),$$

which is readily integrable by quadratures. Knowing $v(x)$ and $u(x)$, we find the solution $y(x)$ of (17.28). ►

Example. Find the general solution of

$$y' + 2xy = xe^{-x^2}.$$

◄ We will seek the solution $y(x)$ of this linear inhomogeneous equation in the form

$$y(x) = u(x)\,v(x).$$

Substituting $y = uv$ into the original equation gives

$$u'v + uv' + 2xuv = xe^{-x^2}$$

or

$$u'v + (v' + 2xv)u = xe^{-x^2}. \tag{17.35'}$$

We define $v(x)$ as a solution of the equation

$$v' + 2xv = 0.$$

By separating the variables, we will get

$$\frac{dv}{v} = -2x\,dx, \quad v = Ce^{-x^2}.$$

We now take any particular solution, e.g., one that corresponds to $C = 1$. Then, by (17.35′), we will obtain

$$e^{-x^2}u' = xe^{-x^2},$$

whence $u' = x$ and $u = x^2/2 + C$.

The general solution of the original equation will be

$$y(x) = u(x)\,v(x) = \left(\frac{x^2}{2} + C\right)e^{-x^2}. \blacktriangleright$$

The method of variation of constants has the advantage that it can be generalized to linear inhomogeneous differential equations of higher orders.

Bernoulli's equation. Some differential equations can be reduced to linear ones by a change of variables. Among such equations is the *Bernoulli equation*

$$\frac{dy}{dx} + p(x)y = q(x)y^{\alpha}, \quad \alpha = \text{const}.$$

At $\alpha = 1$ we obtain the homogeneous linear equation

$$\frac{dy}{dx} + \left[p(x) - q(x)\right]y = 0.$$

At $\alpha = 0$ we have the inhomogeneous linear equation

$$\frac{dy}{dx} + p(x)y = q(x).$$

We suppose, therefore, that $\alpha \neq 0$, $\alpha \neq 1$ (for noninteger α we will assume that $y > 0$).

By substituting $z = y^{-\alpha+1}$ we reduce the Bernoulli equation to a linear equation in $z(x)$.

We can, however, integrate the Bernoulli equation directly by the method of variation of constants. To this end, we proceed as follows. To begin with, we integrate the equation $dy/dx + p(x)y = 0$. Its general solution is $y = Ce^{-\int p(x)\,dx}$. We will look for the solution of the Bernoulli equation in the form

$$y = C(x)\,e^{-\int p(x)\,dx}, \qquad (*)$$

where $C(x)$ is a new unknown function. Substituting this form of $y(x)$ into the Bernoulli equation, we obtain

$$C'(x) = q(x)\,(C(x))^{\alpha}\,e^{(1-\alpha)\int p(x)\,dx}.$$

This is an equation with separable variables in $C(x)$. Integrating gives

$$\frac{(C(x))^{1-\alpha}}{1-\alpha} = \int q(x)e^{(1-\alpha)\int p(x)\,dx}\,dx + C,$$

where C is the constant of integration. From $(*)$ we will then derive the complete integral of the Bernoulli equation

$$y^{1-\alpha}(x) = (1-\alpha)\,e^{(\alpha-1)\int p(x)\,dx}\left[\int q(x)\,e^{(1-\alpha)\int p(x)\,dx}\,dx + C\right].$$

Remark. For $\alpha > 0$ the Bernoulli equation has the obvious solution $y \equiv 0$.

To integrate the Bernoulli equation we can make use of the substitution $y(x) = u(x)\,v(x)$, where $v(x)$ is any nontrivial solution of the equation $v'(x) + p(x)v = 0$, and $u(x)$ is defined as the solution of

$$\frac{du}{dx} = q(x)\,v^{\alpha-1}(x)\,u^{\alpha}.$$

Example. Find the solution of the Bernoulli equation

$$y' - y\tan x = -y^2\cos x.$$

◄ We seek the solution $y(x)$ of the equation in the form $y(x) = u(x)\,v(x)$. Substituting $y = uv$ into the original equation gives

$$u'v + uv' - uv\tan x = -u^2v^2\cos x,$$

or

$$u'v + (v' - v\tan x)\,u = -u^2v^2\cos x.$$

We choose $v(x)$ such that it will be some nonzero solution of the equation

$$v' - v\tan x = 0.$$

We integrate it:

$$\frac{dv}{v} = \frac{\sin x}{\cos x}\,dx, \quad v = \frac{C}{\cos x}.$$

Since we are interested in any particular solution, we put $C = 1$, i.e., take $v = 1/\cos x$. Then for $u(x)$ we will get the equation

$$u' = -u^2.$$

Integration gives $u(x) = 1/(x + C)$.

The general solution $y(x)$ of the original equation is given by

$$y(x) = u(x)\ v(x) = \frac{1}{(x + C)\cos x}. \quad ►$$

Exact differential equations. The equation

$$M(x, y)\,dx + N(x, y)\,dy = 0 \tag{17.36}$$

is said to be an *exact differential equation* if its left-hand side is the total differential of a certain function $u(x, y)$ of two independent variables x and y, i.e.,

$$M(x, y)\,dx + N(x, y)\,dy = du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy.$$

Here $u(x, y) = C$ will be the general integral of (17.36).

We assume that the function $M(x, y)$ and $N(x, y)$ have continuous partial derivatives in y and x, respectively, in a simply connected region D in the xy-plane.

Theorem 17.4. *The necessary and sufficient condition for the left-hand side of* (17.36) *to be the exact* (*total*) *differential of a function $u(x, y)$ of two independent variables x and y is*

$$\frac{\partial M}{\partial y} \equiv \frac{\partial N}{\partial x}. \tag{17.37}$$

◄ *Necessity.* Suppose that the left-hand side of (17.36) is the exact differential of $u(x, y)$, i.e.,

$$M(x, y)\,dx + N(x, y)\,dy = du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy.$$

Then $M = \partial u/\partial x$, $N = \partial u/\partial y$. We differentiate M with respect to y, and N with respect to x

$$\frac{\partial M}{\partial y} = \frac{\partial^2 u}{\partial y\,\partial x}, \qquad \frac{\partial N}{\partial x} = \frac{\partial^2 u}{\partial x\,\partial y}.$$

Since the mixed derivatives are equal,

$$\frac{\partial M}{\partial y} \equiv \frac{\partial N}{\partial x}.$$

This proves the necessity of (17.37).

Sufficiency. Suppose that the condition (17.37) is also sufficient and find $u(x, y)$ such that $du = M(x, y)\,dx + N(x, y)\,dy$, or construct

$$\frac{\partial u}{\partial x} = M(x, y), \quad \frac{\partial u}{\partial y} = N(x, y). \tag{17.38}$$

First of all we find $u(x, y)$ satisfying the first of (17.38). Integrating this with respect to x (assuming y to be constant) gives

$$u = \int M(x, y)\,dx + \varphi(y), \tag{17.39}$$

where $\varphi(y)$ is an arbitrary function of y.

We select $\varphi(y)$ so that the partial derivative of u given by (17.39) with respect to y would be $N(x, y)$. It is always possible to find such a function $\varphi(y)$ subject to (17.37). Indeed, from (17.39),

$$\frac{\partial u}{\partial y} = \frac{\partial}{\partial y}\int M(x, y)\,dx + \varphi'(y).$$

Equating the right-hand side of this to $N(x, y)$ gives

$$\varphi'(y) = N(x, y) - \frac{\partial}{\partial y}\int M(x, y)\,dx. \tag{17.40}$$

The left-hand side of this is independent of x. We will now see that, provided (17.38) is satisfied, its right-hand side does not include x either. With this in mind, we will show that the partial derivative with respect to x of the right-hand side of (17.40) is indentically zero. We thus have

$$\begin{aligned}\frac{\partial}{\partial x}\left[N - \frac{\partial}{\partial y}\int M\,dx\right] &= \frac{\partial N}{\partial x} - \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\int M\,dx\right]\\ &= \frac{\partial N}{\partial x} - \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}\int M\,dx\right],\end{aligned}$$

but

$$\frac{\partial}{\partial x}\int M(x, y)\,dx = M(x, y),$$

hence

$$\frac{\partial}{\partial x}\left[N - \frac{\partial}{\partial y}\int M\,dx\right] = \frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \equiv 0.$$

Now, integrating (17.40) with respect to y, we will get

$$\varphi(y) = \int\left[N - \frac{\partial}{\partial y}\int M\,dx\right]dy + C,$$

where C is the constant of integration. Substituting this into (17.39), we

arrive at the desired function

$$u(x, y) = \int M\,dx + \int\left[N - \frac{\partial}{\partial y}\int M\,dx\right] dy + C,$$

whose exact differential, as is easily verified, is $M(x, y)\,dx + N(x, y)\,dy$. ►

This procedure of constructing $u(x, y)$ is a method of integrating equation (17.36), whose left-hand side is an exact differential.

Example. Check that

$$e^{-y}\,dx - (2y + xe^{-y})\,dy = 0 \qquad (*)$$

is an exact differential equation, and integrate it.

◄ In this case $M = e^{-y}$, $N = -(2y + xe^{-y})$

$$\frac{\partial M}{\partial y} = -e^{-y}, \qquad \frac{\partial N}{\partial x} = -e^{-y}, \quad \text{hence} \quad \frac{\partial M}{\partial y} \equiv \frac{\partial N}{\partial x};$$

and so $(*)$ is an exact differential equation.

We now want to find u (see (17.39)):

$$u = \int M(x, y)\,dx + \varphi(y) = \int e^{-y}\,dx + \varphi(y)$$

or

$$u = xe^{-y} + \varphi(y). \qquad (**).$$

Finding $\partial u/\partial y$ from $(**)$ and equating $\partial u/\partial y$ to $N(x, y) = -2y - xe^{-y}$, we obtain

$$-xe^{-y} + \varphi'(y) = -2y - xe^{-y}.$$

Thus, $\varphi'(y) = -2y$, and hence

$$\varphi(y) = -y^2 + C, \qquad C = \text{const}.$$

Substituting this into $(**)$ gives

$$u = xe^{-y} - y^2 + C; \quad xe^{-y} - y^2 = C,$$

i.e., the general integral of the original equation. ►

It is sometimes possible to find a function $\mu(x, y)$ such that $\mu M\,dx + \mu N\,dy$ will be an exact differential, although $M\,dx + N\,dy$ may not be one. Such a function $\mu(x, y)$ is called an *integrating factor.*

It can be shown that for the first-order equation $M(x, y)\,dx + N(x, y) = 0$ with $M(x, y)$ and $N(x, y)$ subject to certain conditions there always exists an integrating factor, but to deduce it from the condition $\frac{\partial(\mu M)}{\partial y} = \frac{\partial(\mu N)}{\partial x}$ in the general case means to integrate a partial differential equation, which as a rule is a more difficult task.

Problem. Find the integrating factor for the linear differential equation

$$\frac{dy}{dx} + p(x)y = q(x).$$

Hint: Seek the factor in the form $\mu = \mu(x)$.

17.5 Riccati Equation

The equation

$$\frac{dy}{dx} = q(x) + p(x)y + r(x)y^2, \tag{17.41}$$

where $q(x)$, $p(x)$, and $r(x)$ are known functions, is called the *Riccati equation*. If p, q, and r are constants, then it is integrated by separating the variables

$$\int \frac{dy}{q + py + ry^2} = x + C.$$

When $r(x) \equiv 0$, equation (17.41) is linear, and when $q(x) \equiv 0$, it is the Bernoulli equation. In the general case, equation (17.41) is not integrable by quadratures.

We will now discuss some properties of the Riccati equation.

Theorem 17.5. *Given one particular solution of the Riccati equation, its general solution can be found by quadratures.*

◄ Suppose we know the particular solution $y = y_1(x)$ of (17.41), then

$$y_1'(x) \equiv q(x) + p(x)y_1(x) + r(x)y_1^2(x). \tag{17.42}$$

Putting $y = y_1(x) + z(x)$, where $z(x)$ is a new desired function, we obtain, by (17.42),

$$\frac{dz}{dx} - (p(x) + 2r(x)y_1(x))z = r(x)z^2.$$

~~This is~~ the Bernoulli equation, which is integrated by quadratures. ►

Example. Integrate the Riccati equation

$$y' - y^2 + 2e^x y = e^{2x} + e^x,$$

if we know its particular solution $y_1 = e^x$.

◄ Putting $y = e^x + z$, we will have for $z(x)$

$$\frac{dz}{dx} = z^2, \quad \text{hence} \quad z = \frac{1}{C - x}.$$

The solution of the original equation will be

$$y(x) = e^x + \frac{1}{C - x}. \quad ►$$

A special case of (17.41) is the *special* Riccati equation

$$\frac{dy}{dx} + ay^2 = bx^\alpha \qquad (x > 0), \tag{17.43}$$

where a, b, and α are constants.

At $\alpha = 0$ we have $dy/dx = b - ay^2$ and the equation is integrated by separation of variables.

At $\alpha = -2$ we get $dy/dx + ay^2 = b/x^2$. Setting $y = 1/z$, where z is a new unknown function, we get

$$-\frac{1}{z^2}\frac{dz}{dx} + \frac{a}{z^2} = \frac{b}{x^2} \quad \text{hence} \quad \frac{dz}{dx} = a - b\left(\frac{z}{x}\right)^2.$$

This equation is homogeneous in x and z; it is integrated by quadratures.

Along with $\alpha = 0$ and $\alpha = -2$ there exist an infinite variety of other values of α at which the Riccati equation (17.43) is integrated by quadratures. They are given by

$$\alpha = \frac{4k}{-2k+1} \qquad (k = \pm 1, \pm 2, \ldots).$$

For all other values of α the solution of the Riccati equation (17.43) is not expressed by quadratures. If then we set in it $y = u'/(au)$, where $u = u(x)$ is a new unknown function, we will then arrive at the second-order equation $d^2u/dx^2 - abx^\alpha u = 0$, whose solution can be expressed in terms of the Bessel functions (see Chap. 18).

17.6 Differential Equations Insolvable for the Derivative

Consider the general case of the first-order equation

$$F(x, y, y') = 0 \tag{17.44}$$

which is insolvable for the derivative. It has some distinctive features. For instance, the equation $(y')^2 + 1 = 0$ has no real-valued solutions at all. The solutions of the equation $(y')^2 = 1$ are the straight lines $y = \pm x + C$, so that through each point in the xy-plane pass two mutually perpendicular integral lines. The field of integral curves of the equation $(y')^2 = 1$ is obtained by superposition of the field for $y' = 1$ and $y' = -1$. If (17.44) can be solved for y', we obtain equations of the form $y' = f_i(x, y)$, which can sometimes be integrated by the methods discussed above.

We now introduce the concept of the *general solution* (general integral) for (17.44). Suppose that in the neighbourhood of a point (x_0, y_0) we can solve this equation for the derivative, i.e., we have the equations

$$y' = f_i(x, y) \qquad (i = 1, 2, \ldots, m)$$

and suppose further that each of these equations has the general solution

$$y = \varphi_i(x, C) \qquad (i = 1, 2, \ldots, m) \tag{17.45}$$

or the general integral

$$\Psi_i(x, y, C) = 0 \qquad (i = 1, 2, \ldots, m). \tag{17.46}$$

The multitude of the general solutions (17.45) (or the integrals (17.46)) is known as the *general solution* (*integral*) of (17.44). Thus, the equation $(y')^2 = 1$ breaks down into two: $y' = 1$, $y' = -1$; their general solutions $y = x + C$, $y = -x + C$ in combination constitute the general solution of the original equation $(y')^2 = 1$. The general integral of this equation is often written as $(y - x - C)(y + x - C) = 0$.

It is not always easy, however, to solve (17.44) for y' and the resultant equation $y' = f_i(x, y)$ often cannot be integrated by quadratures. We will now look at some of the procedures of integrating (17.44).

(1) Let equation (17.44) have the form

$$F(y') = 0. \tag{17.47}$$

Suppose that there exists at least one real-valued solution $y' = k_i$ of this equation. Since this equation does not contain x and y, then k_i is a constant. Integrating the equation $y' = k_i$ gives

$$y = k_i x + C, \quad \text{or} \quad k_i = \frac{y - C}{x}.$$

But k_i is a root of $F(k) = 0$, accordingly, $F((y - C)/x) = 0$ is the integral of the equation.

For example the equation $(y')^7 - (y')^5 + y' - 3 = 0$ has the integral

$$\left(\frac{y - C}{x}\right)^7 - \left(\frac{y - C}{x}\right)^5 + \frac{y - C}{x} - 3 = 0.$$

(2) Equation (17.44) may have the form

$$F(y, y') = 0. \tag{17.48}$$

If it is difficult to solve this equation for y', then it may be advisable to introduce a parameter t and to replace (17.48) by the two equations

$$y = \varphi(t), \quad y' = \psi(t), \quad (t_0 \leqslant t \leqslant t_1),$$

such that $F(\varphi(t), \psi(t)) \equiv 0$, $t \in (t_0, t_1)$. Since $dy = y'dx$, then

$$dx = \frac{dy}{y'} = \frac{\varphi'(t)}{\psi(t)}\, dt, \quad \text{hence} \quad x = \int \frac{\varphi'(t)}{\psi(t)}\, dt + C.$$

The desired integral curves are thus determined by the equations in para-

metric form

$$x = \int \frac{\varphi'(t)}{\psi(t)}\, dt + C, \quad y = \varphi(t).$$

Example. Integrate the equation $y^{2/3} + (y')^{2/3} = 1$.

◄ We put $y = \cos^3 t$, $y' = \sin^3 t$, then

$$dx = \frac{dy}{y'} = \frac{-3\cos^2 t \sin t}{\sin^3 t}\, dt = -3\frac{\cos^2 t}{\sin^2 t}\, dt.$$

Further, we find

$$x = -3\int \frac{\cos^2 t}{\sin^2 t}\, dt = 3t + 3\cot t + C$$

and the parametric equations of the desired integral curves

$$x = 3t + 3\cot t + C, \quad y = \cos^3 t. \text{ ►}$$

If equation (17.48) is readily solvable for y, the parameter is normally y'. In fact, if $y = \varphi(y')$, then, putting $y' = p$, we get $y = \varphi(p)$, so that

$$dx = \frac{dy}{y'} = \frac{\varphi'(p)}{p}\, dp; \quad x = \int \frac{\varphi'(p)}{p}\, dp + C.$$

The parametric equation of the integral curves will be

$$x = \int \frac{\varphi'(p)}{p}\, dp + C, \quad y = \varphi(p).$$

Excluding p, we obtain the general integral $\Phi(x, y, C) = 0$ of (17.48).

Example. Integrate the equation $y\sqrt{y' - 1} = 2 - y'$.

◄ We solve the equation for y to obtain

$$y = \frac{2 - y'}{\sqrt{y' - 1}}.$$

We then set $y' = p$, and find

$$y = \frac{2 - p}{\sqrt{p - 1}}.$$

Further,

$$dx = \frac{dy}{y'} = -\frac{dp}{2(p - 1)^{3/2}}; \quad x = \frac{1}{\sqrt{p - 1}} + C.$$

We thus find the parametric equations of the integral curves

$$x = \frac{1}{\sqrt{p - 1}} + C, \quad y = \frac{2 - p}{\sqrt{p - 1}}.$$

We can easily exclude p. Really, from the first equation of the above system we find

$$\frac{1}{\sqrt{p-1}} = x - C, \quad p = \frac{1}{(x - C)^2} + 1.$$

If we substitute p into the second equation, we will have the general solution of the original equation: $y = x - C - 1/(x - C)$. ►

(3) Let (17.44) have the form

$$F(x, y') = 0. \tag{17.49}$$

If this equation is hard to solve for y', then, as in the previous case, it is expedient to introduce the parameter t and to replace (17.49) by two equations:

$$x = \varphi(t), \quad y' = \psi(t) \quad (t_0 \leqslant t \leqslant t_1).$$

Then

$$dy = y'\,dx = \psi(t)\,\varphi'(t)\,dt,$$
$$y = \int \psi(t)\,\varphi'(t)\,dt + C.$$

The integral curves of (17.49) are thus given in parametric form by

$$x = \varphi(t), \quad y = \int \psi(t)\,\varphi'(t)\,dt + C.$$

If equation (17.49) can easily be solved for x, $x = \varphi(y')$, it is convenient to take as the parameter $y' = p$. Then $x = \varphi(p)$ and $dy = y'\,dx = p\,\varphi'(p)\,dp$, whence $y = \int p\,\varphi'(p)\,dp + C$.

Example. Solve the equation $x = (y')^3 - y' - 1$.

◄ Put $y' = p$, then $x = p^3 - p + 1$. Further,

$$dy = y'\,dx = p\,(3p^2 - 1)\,dp; \quad y = \frac{3p^4}{4} - \frac{p^2}{2} + C.$$

In parametric form, the family of integral curves of this equation is given by

$$x = p^3 - p - 1, \qquad y = \frac{3p^4}{4} - \frac{p^2}{2} + C. \ ►$$

(4) *Lagrange equation.* A differential equation of the type

$$y = x\,\varphi(y') + \psi(y') \tag{17.50}$$

linear in x and y, is known as the Lagrange equation. Here φ and ψ are known functions.

Introducing the parameters $dy/dx = p$ gives

$$y = x\,\varphi(p) + \psi(p). \tag{17.51}$$

This formula relates x, y and p. To derive a second relation for determining x and y as functions of p, we will differentiate (17.51) with respect to x

$$p = \varphi(p) + x\varphi'(p)\frac{dp}{dx} + \psi'(p)\frac{dp}{dx},$$

whence

$$p - \varphi(p) = \left[x\varphi'(p) + \psi'(p)\right]\frac{dp}{dx} \tag{17.52}$$

or

$$\left[p - \varphi(p)\right]\frac{dx}{dp} = x\varphi'(p) + \psi'(p). \tag{17.53}$$

Equation (17.53) is linear in x and dx/dp, and so it can readily be integrated, e. g., by variation of constants. We now derive the general solution $x = \omega(p, C)$ of (17.53) and add to it the equation $y = x\varphi(p) + \psi(p)$, to obtain the parametric equations of the desired integral curves.

In passing from (17.52) to (17.53) we divided by dp/dx. This causes solutions for which p is constant to be lost and so $dp/dx = 0$. Taking p to be constant, we note that (17.52) is only satisfied when p is a root of the equation $p - \varphi(p) = 0$. Thus, if the equation $p - \varphi(p) = 0$ has real roots $p = p_i$, then to the solutions of the Lagrange equation we have found above we have to add the solutions

$$\begin{cases} y = x\varphi(p) + \psi(p), \\ p = p_i, \end{cases} \quad \text{or} \quad y = x\varphi(p_i) + \psi(p_i),$$

which are straight lines.

(5) *Clairaut's equation.* An equation of the form

$$y = xy' + \psi(y') \tag{17.54}$$

is called the Clairaut equation.

Putting $y' = p$, we obtain $y = xp + \psi(p)$. Differentiating with respect to x gives

$$p = p + x\frac{dp}{dx} + \psi'(p)\frac{dp}{dx} \quad \text{or} \quad \left[x + \psi'(p)\right]\frac{dp}{dx} = 0.$$

Therefore, either $dp/dx = 0$, and hence $p = C$, or $x + \psi'(p) = 0$. In the first case, excluding p, we will find the family of curves

$$y = Cx + \psi(C),$$

i.e., the general solution of the Clairaut equation. It is found without quad-

ratures and describes a one-parameter family of lines. In the case of $x + \psi'(p) = 0$ the solution is given

$$y = xp + \psi(p), \quad x = -\psi'(p). \tag{17.55}$$

It can be shown that, as a rule, the integral curve (17.55) is the envelope of the family obtained.

Example. Solve the Clairaut equation $y = xy' - (y')^2$.

◀ The general solution of this equation is immediate: $y = Cx - C^2$. Another (singular) solution is given by

$$y = xp - p^2, \quad x = 2p.$$

Excluding the parameter p, we find $y = x^2/4$, the envelope of the lines $y = Cx - C^2$ (Fig. 17.7). ▶

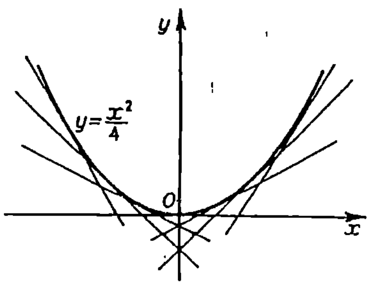


Fig. 17.7

Equations of the form $F(x, y, y') = 0$, generally speaking, have several integral curves passing through a certain point (x_0, y_0). Indeed, solving the equation for y', we as a rule obtain several real-valued functions $y' = f_i(x, y)$, $i = 1, 2, \ldots, m$. If every $y' = f_i(x, y)$ in the neighbourhood of (x_0, y_0) satisfies the conditions of Theorem 17.1 on the existence and uniqueness of the solution, then for each of these equations there exists a unique solution meeting the condition $y(x_0) = y_0$. Therefore, speaking about the uniqueness of the solution to $F(x, y, y') = 0$ subject to the condition $y(x_0) = y_0$, we generally mean that through a *given* point (x_0, y_0) in a *given* direction passes only one integral curve of the original equation $F(x, y, y') = 0$.

For example, for the solutions to the equation $(y')^2 - 1 = 0$ the uniqueness property in this sense is satisfied everywhere, since through each point (x_0, y_0) in the xy-plane pass two integral curves but in different directions. For the Clairaut equation $y = xy' - (y')^2$ (see the above example) through the point (0, 0) also pass two integral lines: the straight line $y = 0$, which

enters into the general solution of this equation, and the parabola $y = x^2/4$. These lines have at (0, 0) the *same* direction since $y' = 0$. Therefore, at (0, 0) the uniqueness property is violated.

Theorem 17.6. *In a certain neighbourhood of the point* (x_0, y_0, y_0'), *where* y_0' *is one of the real-valued solutions of the equation* $F(x_0, y_0, y') = 0$, *the function* $F(x, y, y')$ *satisfies the conditions*:

(1) *it is continuous in all its arguments*;

(2) *there exists a nonzero derivative* $\partial F/\partial y'$;

(3) *there exists a bounded derivative* $\partial F/\partial y$: $|\partial F/\partial y| \leqslant N$.

Then there exists an interval $[x_0 - h_0, x_0 + h_0]$ *on which there exists a unique solution* $y = y(x)$ *of the equation* $F(x, y, y') = 0$ *satisfying the condition* $y(x_0) = y_0$ *and for which* $y'(x_0) = y_0'$.

17.7 Geometrical Aspects of First-Order Differential Equations. Orthogonal Trajectories

The general solution $y = \varphi(x, C)$ of a differential equation of the first order defines in a plane a family of curves dependent only on one parameter C.

Let us now state a problem that is in a sense inverse: given a one-parameter family of curves $y = \varphi(x, C)$, construct a differential equation for which $y = \varphi(x, C)$ will be the general solution.

We thus have

$$y = \varphi(x, C), \tag{17.56}$$

where C is a parameter. Differentiating (17.56) with respect to x gives

$$y' = \varphi_x'(x, C). \tag{17.57}$$

If the right-hand side of (17.57) no longer contains C, then (17.57) will be the differential equation of the family of curves (17.56). For example, if $y = x + C$, then $y' = 1$ will be the differential equation of the family $y = x + C$. If now the right-hand side of (17.57) contains C, we can solve (17.56) for C as a function of x and y to find

$$C = \psi(x, y). \tag{17.58}$$

Substituting this into (17.57), we will obtain the first-order differential equation

$$y' = \varphi'(x, \psi(x, y)). \tag{17.59}$$

It is easily seen that $y = \varphi(x, C)$ is the solution of (17.59) at any C.

If $x, y,$ and C are related by

$$\Phi(x, y, C) = 0, \tag{17.60}$$

then differentiating this with respect to x gives

$$\Phi'_x + \Phi'_y \cdot y' = 0. \tag{17.61}$$

Excluding C from (17.60) and (17.61) gives

$$F(x, y, y') = 0. \tag{17.62}$$

It can be shown that (17.60) is the general integral of (17.62).

Orthogonal trajectories. A number of applications lead to the following problem. Given a family of curves $\Phi(x, y, C) = 0$, find a family of curves $\Psi(x, y, C) = 0$ such that each curve of $\Phi(x, y, C) = 0$ passing through a

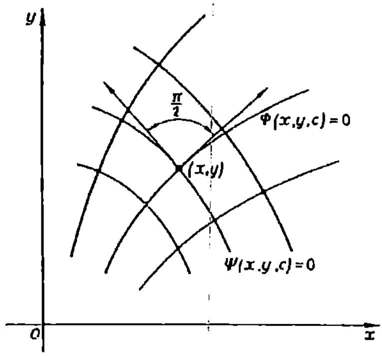


Fig. 17.8

point (x, y) would meet a curve of $\Psi(x, y, C) = 0$ at a right angle, i.e., such that tangents to the curve of $\Phi = 0$ and $\Psi = 0$ at (x, y) would be orthogonal (Fig. 17.8). The family $\Psi(x, y, C) = 0$ is called a *family of orthogonal trajectories* to $\Phi(x, y, C) = 0$, and vice versa. If, for instance, the curves of $\Phi = 0$ are the lines of force of a force field, then orthogonal trajectories are equipotential lines.

Analytically, this means the following. If $F(x, y, y') = 0$ is the differential equation of the family $\Phi(x, y, C) = 0$, then the differential equation of trajectories orthogonal to $\Phi = 0$ has the form

$$F\left(x, y, -\frac{1}{y'}\right) = 0$$

(the slopes of the tangents to the curves $\Phi = 0$ and $\Psi = 0$ at every point must be related by the orthogonality condition $k_1 k_2 = -1$).

To sum up: if we want to find orthogonal trajectories to a family $\Phi(x, y, C) = 0$, we have to derive the differential equation $F(x, y, y')$ of the family and replace in it y' by $-1/y'$. Integrating the resultant equation will give the family of orthogonal trajectories.

Example. Find the orthogonal trajectories of the family

$$x^2 + y^2 = C^2 \tag{17.63}$$

of circles with centre at the origin.

◄ We will set up the differential equation of the family (17.63). Differentiating (17.63) with respect to x gives $2x + 2yy' = 0$ or $x + yy' = 0$. Hence

$$y' = -\frac{x}{y}.$$

This is the differential equation of the family. If now we replace in it y' by $-1/y'$, we will find the differential equation of the family of orthogonal trajectories

$$-\frac{1}{y'} = -\frac{x}{y} \quad \text{or} \quad y' = \frac{y}{x}.$$

Integrating the last equation, we find that the desired orthogonal trajectories are (Fig. 17.9)

$$\begin{aligned} y &= Cx \quad (x \neq 0), \\ x &= 0 \quad (y \neq 0). \end{aligned}$$ ►

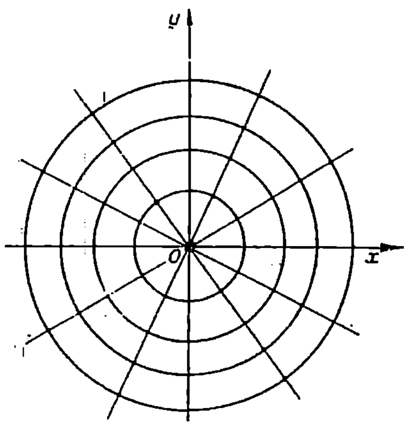


Fig. 17.9

Exercises

Use the method of isoclines to construct the integral curves of the following differential equations:

1. $y' = x + y$. 2. $y' = (y - 1)x$.

By successive approximations solve the Cauchy problem:

3. $y' = 2xy$, $y(0) = 1$.

Integrate the differential equations:

4. $2xy\,dy + (1 + y^2)\,dx = 0$. 5. $e^{3y}\sin^2 x\,dx + \cos^2 x\,dy = 0$.

6. $x\dfrac{dy}{dx} = y + xe^{y/x}$, $y(1) = \ln 2$. 7. $xy' = y + \sqrt{y^2 - x^2}$.

8. $y' = \dfrac{x + y - 2}{y - x - 4}$. 9. $y' + 2xy = xe^{-x^2}$.

10. $y' - y\cot x = 2x\sin x$, $y(\pi/2) = 1$.

11. $x^2y' + 2xy = e^x$, $y(1) = 1$. 12. $y' = \dfrac{1}{2x - y^2}$.

13. $xy' + y = y^2\ln x$. 14. $xyy' = 2y^2 - 3x^2$.

15. $(xy + x^2y^3)\,dy = dx$. 16. $x^2yy' + y^2x = 1$.

17. $\left(2x - 1 - \dfrac{y}{x^2}\right)dx - \left(2y - \dfrac{1}{x}\right)dy = 0$.

18. $\left(y^2 + \dfrac{y}{\cos^2 x}\right)dx + (2xy + \tan x)\,dy = 0$. 19. $2y = xy' + y'\ln y'$.

20. Find the orthogonal trajectories of the family of hyperbolas $x^2 - y^2 = a^2$, where a is a numerical parameter.

Answers

1.

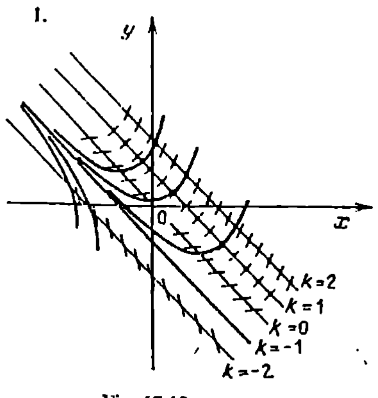


Fig. 17.10

2.

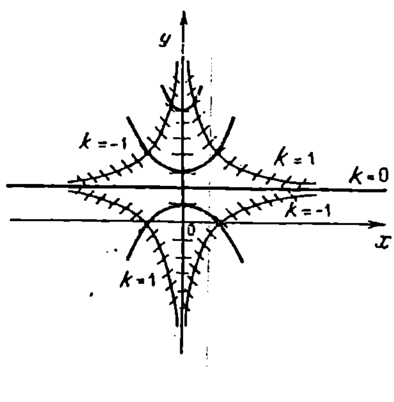


Fig. 17.11

3. $y = e^{x^2}$. 4. $x + 1 + y^2) = C$. 5. $\tan x - x - e^{-3y}/3 = C$. 6. $e^{-y+x} = \ln \dfrac{\sqrt{e}}{x}$.
7. $Cx^2 = y + \sqrt{y^2 - x^2}$. 8. $x^2 + 2xy - y^2 - 4x + 8y = C$.
9. $y = \left(\dfrac{x^2}{2} + C\right) e^{-x^2}$. 10. $y = (x^2 + 1 - \pi^2/4)\sin x$.
11. $y = (1 - e + e^y)/x^2$. 12. $x = Ce^{2y} + \dfrac{y^2}{2} + \dfrac{y}{2} + \dfrac{1}{4}$. 13. $y = 1/(Cx + \ln x + 1)$.
14. $\sqrt{y^2 - 3x^2} = Cx^2$. 15. $x = 1/(Ce^{-y^2+2} + y^2 - 2)$. 16. $y^2 = (C + x)/x^2$.
17. $x^2 - x + y/x - y^2 = C$. 18. $xy^2 + y\tan x = C$.
19. $\begin{cases} x = Cp - \ln p - 2, \\ y = Cp^2/2 - p. \end{cases}$ 20. $xy = C$.

Chapter 18

Higher-Order Differential Equations

18.1 Cauchy Problem

Let

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) \tag{18.1}$$

be a differential equation of nth order, solved for the highest-order derivative $y^{(n)}$.

What conditions have to be specified so that we could obtain a definite particular solution of equation (18.1)? For a first-order differential equation $y' = f(x, y)$ it is sufficient to specify the value y_0 of a particular solution at some value x_0 of the independent variable x, i.e., to specify a point (x_0, y_0) through which the integral curve of the equation must pass. For higher-order equations this is no longer sufficient. For example, the equation $y'' = 0$ has the solutions $y = C_1x + C_2$, where C_1 and C_2 are arbitrary constants. The equation $y = C_1x + C_2$ defines a two-parameter family of straight lines in the xy-plane. To isolate a straight line it is insufficient to know the point (x_0, y_0) through which the line must pass, since we also need the slope $y'|_{x=x_0} = y_0'$ of the straight line.

In the general case of the nth-order differential equation (18.1) to obtain a particular solution we should specify n conditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \quad \dots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}, \tag{18.2}$$

where $y_0, y_0', \dots, y_0^{(n-1)}$ are some numbers. The collection of these conditions is known as the *initial conditions for the differential equation* (18.1).

The Cauchy problem for this equation is formulated as follows: find the solution of the differential equation (18.1) satisfying the initial conditions (18.2).

The existence and uniqueness of the solution of the Cauchy problem is taken care of by

Theorem 18.1. *Let*

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$$

be a differential equation of nth order solved for the highest-order derivative.

If the right-hand side of (18.1) *is continuous as a function of* $n+1$ *arguments* $x, y, y', \ldots, y^{(n-1)}$ *in a certain neighbourhood* Ω *of the point* $M_0(x_0, y_0, y_0', \ldots, y_0^{(n-1)})$ (in Fig. 18.1 for $n = 2$), *then there exists an interval* $x_0 - h_0 < x < x_0 + h_0$ *on the x-axis on which there exists at least one solution* $y = \varphi(x)$ *of* (18.1) *satisfying the initial conditions*

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \ \ldots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}.$$

Further, if the function $f(x, y, y', \ldots, y^{(n-1)})$ *has bounded partial derivatives* $\dfrac{\partial f}{\partial y}, \dfrac{\partial f}{\partial y'}, \ldots, \dfrac{\partial f}{\partial y^{(n-1)}}$ *in* Ω *then such a solution is unique.*

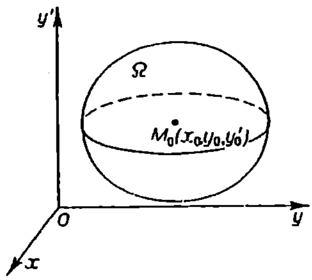


Fig. 18.1

For example, for the equation $y'' = e^{-x^2}y + \sin y'$ the right-hand side $f = e^{-x^2}y + \sin y'$, viewed as a function of three independent variables, x, y, y', is continuous everywhere and has derivatives $\dfrac{\partial f}{\partial y} = e^{-x^2}$, and $\dfrac{\partial f}{\partial y'} = \cos y'$, which are bounded everywhere. Therefore, whatever the numbers x_0, y_0, and y_0', there exists a unique solution of the equation such that it obeys the initial conditions $y|_{x=x_0} = y_0$ and $y'|_{x=x_0} = y_0'$.

Definition. The *general solution* of equation (18.1) in a certain domain Ω where there exists a unique solution of the Cauchy problem is an *n-parametric family S of functions* $y = \varphi(x, C_1, C_2, \ldots, C_n)$ dependent on x and n arbitrary constants $C_1, C_2, \ldots, C_n$, such that

(1) for every admissible $C_1, C_2, \ldots, C_n$ the function $y = \varphi(x, C_1, C_2, \ldots, C_n) \in S$ is a solution of (18.1), i.e.,

$$\varphi^{(n)}(x, C_1, C_2, \ldots, C_n) \equiv f(x, \varphi(x, C_1, C_2, \ldots, C_n), \ldots, \varphi^{(n-1)}(x, C_1, C_2, \ldots, C_n)),$$

$x \in (x_0 - h, x_0 + h)$;

(2) for any initial conditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \ \dots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}$$

such that the point $(x_0, y_0, y_0', \dots, y_0^{(n-1)})$ belongs to Ω where there exists a unique solution of the Cauchy problem for (18.1), we can find C_1^0, C_2^0, ..., C_n^0 such that the solution $y = \varphi(x, C_1^0, C_2^0, \dots, C_n^0) \in S$ satisfies the initial conditions.

A solution derived from the general one for specific values of C_1, C_2, ..., C_n is called a *particular solution*. Its plot, a curve in the xy-plane, is called the *integral curve* of the equation.

Relation $\Phi(x, y, C_1, C_2, \dots, C_n) = 0$, which implicitly defines the general solution, is known as the *general* or *complete integral* of (18.1).

Problem. Show that the function $y = C_1 \cos x + C_2 \sin x$ is the general solution of the equation $y'' + y = 0$.

18.2 Reducing the Order of Higher-Order Equations

(1) An equation of the form

$$y^{(n)} = f(x), \tag{18.3}$$

where $f(x)$ is a known continuous function, is integrable by quadratures. Really, considering that $y^{(n)} = (y^{(n-1)})'$ and integrating with respect to x both sides of the equation, we will have

$$y^{(n-1)} = \int f(x)\,dx + C_1,$$

i.e., we arrive at an equation of the same form as the original one. We further find

$$y^{(n-2)} = \int \left(\int f(x)\,dx\right)dx + C_1 x + C_2.$$

In n steps we will obtain the general solution of (18.1)

$$y(x) = \int\left[\int\left(\dots\int f(x)\,dx\right)dx\right]dx$$

$$+ C_1 \frac{x^{n-1}}{(n-1)!} + C_2 \frac{x^{n-2}}{(n-2)!} + \dots + C_n.$$

Example. Find the general solution of the equation $y'' = 2x$.

◄ Integrating twice, we get the desired general solution

$$y' = \int 2x\,dx = x^2 + C_1, \quad y = \frac{x^3}{3} + C_1 x + C_2. \ ►$$

(2) If an equation does not contain the desired function and its derivatives through order $k-1$, i.e., it has the form

$$F(x, y^{(k)}, y^{(k+1)}, \ldots, y^{(n)}) = 0, \tag{18.4}$$

then the order of the equation can be reduced down to $n-k$ by the change $y^{(k)} = p(x)$. The equation becomes

$$F(x, p, p', \ldots, p^{(n-k)}) = 0.$$

Suppose we can integrate this equation to obtain

$$p = \Psi(x, C_1, C_2, \ldots, C_{n-k}).$$

Notice that $p = y^{(k)}(x)$, then

$$y^{(k)} = \Psi(x, C_1, C_2, \ldots, C_{n-k}).$$

From this equation we find $y(x)$ by integrating k times.

Example. Find the general solution of the equation $y''' - y''/x = 0$.

◄ Put $y'' = p(x)$, then $y''' = p'(x)$, and the equation will become

$$\frac{dp}{dx} - \frac{p}{x} = 0.$$

Separating the variables in this equation, we will find $p = C_1x$ or $y'' = C_1x$. From this we readily find the general solution

$$y = \frac{C_1}{6}x^3 + C_2x + C_3. \text{ ►}$$

(3) If a differential equation does not explicitly contain an independent variable x, i.e., it has the form

$$F(y, y', \ldots, y^{(n)}) = 0, \tag{18.5}$$

the order of this equation can be reduced by one by substituting $y' = p(y)$, where $p = p(y)$ is treated as a new unknown function, and y is taken to be an independent variable. We will have then to express all the derivatives d^ky/dx^k, $k = 1, 2, \ldots, n$, in terms of derivatives of p with respect to y:

$$\frac{dy}{dx} = p(y), \quad \frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dy}(p(y))\frac{dy}{dx} = p\frac{dp}{dy},$$

$$\frac{d^3y}{dx^3} = \frac{d}{dx}\left(\frac{d^2y}{dx^2}\right) = \frac{d}{dy}\left(p\frac{dp}{dy}\right)\frac{dy}{dx} = p\left(\frac{dp}{dy}\right)^2 + p^2\frac{d^2p}{dy^2}$$

and so on.

We see that any derivative d^ky/dx^k, $k = 1, 2, \ldots, n$, is expressed through the derivatives of p with respect to y of the order not higher than $k-1$, i.e., the order of the equation has been reduced by one.

Example. Integrate the equation

$$yy'' + (y')^2 = 0 \tag{*}$$

◄ We put $y' = p(y)$, then $y'' = p(dp/dy)$ and the equation becomes

$$yp\frac{dp}{dy} + p^2 = 0.$$

We now reduce this by p, $p \neq 0$, and separate the variables. We thus get

$$p(y) = \frac{\tilde{C}_1}{y} \quad \text{or} \quad \frac{dy}{dx} = \frac{\tilde{C}_1}{y},$$

hence

$$y^2/2 = \tilde{C}_1 x + \tilde{C}_2 \quad \text{or} \quad y = \sqrt{C_1 x + C_2}. \tag{**}$$

The case of $p = 0$ gives the solution $y = C$, which is included in (**). ►

It always pays to see whether or not the left-hand side of the equation is a total differential of some expression. So, equation (*) can be written as

$$\frac{d}{dx}(yy') = 0,$$

whence

$$yy' = \frac{C_1}{2}; \quad \frac{d}{dx}\left(\frac{y^2}{2}\right) = \frac{C_1}{2}; \quad y^2 = C_1 x + C_2.$$

The common equation $y'' = f(y)$ can readily be integrated by quadratures, if we multiply its both sides by y'. We offer the reader to do this.

Remark 1. Consider the second-order equation

$$y'' + p_1(x)y' + p_2(x)y = 0, \tag{18.6}$$

which is linear in the desired function $y(x)$ and its derivatives y' and y''. We put

$$y(x) = u(x)\,v(x), \tag{18.7}$$

where $u(x)$ and $v(x)$ are new functions, one of which can be selected in an arbitrary manner. Substituting $y(x)$ in the form (18.7) into (18.6) we obtain for $u(x)$

$$u'' + \left(2\frac{v'}{v} + p_1\right)u' + \left(\frac{v'' + p_1 v' + p_2 v}{v}\right)u = 0. \tag{18.8}$$

If we know one solution $y_1(x) \neq 0$ of (18.6), then we can take $v = y_1(x)$. In (18.8) the term containing $u(x)$ will then vanish (if $v = y_1(x)$, then $v'' + p_1 v' + p_2 v \equiv 0$, since $y_1(x)$ is assumed to be a solution of (18.6)).

Equation (18.8) will then become

$$u'' + \left(2\frac{y_1'}{y_1} + p_1\right)u' = 0.$$

It can readily be integrated. As a result, we will find the general solution of (18.6).

If we put

$$v(x) = e^{-\frac{1}{2}\int p_1(x)\,dx}, \tag{18.9}$$

then in (18.8) the term with the first derivative will vanish, and the equation will assume the form $u'' + q(x)\,u = 0$. This transformation is useful for qualitative examination of the equation and for approximate methods of solving the equation.

Consider, for example, the Bessel differential equation

$$x^2y'' + xy' + (x^2 - \nu^2)\,y = 0, \tag{18.10}$$

where ν is a numerical parameter. The solutions of the equation are the Bessel functions, which play an important role in many problems of physics. We represent (18.10) in the form

$$y'' + \frac{1}{x}\cdot y' + \left(1 - \frac{\nu^2}{x^2}\right)y = 0.$$

Here $p_1(x) = 1/x$ and by (18.9) we have

$$v(x) = e^{-\frac{1}{2}\int\frac{dx}{x}} = x^{-\frac{1}{2}}, \quad x > 0.$$

Putting $y(x) = x^{-1/2}u(x)$, we obtain for $u(x)$:

$$u'' + \left(1 - \frac{\nu^2 - 1/4}{x^2}\right)u = 0.$$

This equation is convenient to examine the behaviour of the Bessel functions at large x.

Remark 2. When solving the Cauchy problem for higher-order equations it may be a good idea to find the values of C_i in the course of solution, and not after the general solution of the equation has been found. This is because the integration is sometimes simplified dramatically, when C_i assume concrete numerical values, whereas with arbitrary C_i integration in elementary functions is more difficult, if not impossible.

Consider, for example, the following Cauchy problem:

$$y'' = 2y^3, \quad y|_{x=0} = 1, \quad y'|_{x=0} = 1.$$

Letting $y' = p(y)$, we obtain $p(dp/dy) = 2y^3$, whence $p^2 = y^4 + C_1$ or $dy/dx = \sqrt{y^4 + C_1}$. Separating the variables, we will find $x + C_2 = \int (y^4 + C_1)^{-1/2}\, dy$. In the right-hand side of the last identity we have an integral of a differential binomial. Here $m = 0$, $n = 4$, $p = -1/2$, so that this integral cannot be expressed through a finite combination of elementary functions. However, if we turn to the initial conditions, then we will have $C_1 = 0$. This immediately gives $dy/dx = y^2$, and so from the initial conditions we will get

$$y = \frac{1}{1 - x}.$$

Problem. Find two solutions of the Cauchy problem for the equation $y'' = 3\sqrt[3]{y'^2}$ with the initial conditions $y(0) = y'(0) = 0$. One may wonder, whether or not this fact is at variance with Theorem 17.1 (on existence and uniqueness of the Cauchy problem)?

18.3 Linear Homogeneous Differential Equations of Order n

A *linear differential equation of nth order* is an equation linear in some unknown function and all its derivatives. It has the form

$$a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_n(x)y = g(x),$$

where $a_0(x)$, $a_1(x)$, ..., $a_n(x)$, $g(x)$ are functions defined on an interval (α, β). If $g(x) \equiv 0$ on the interval, then the equation is called a linear *homogeneous* equation, otherwise the equation is called *inhomogeneous*.

Suppose we have a linear homogeneous differential equation

$$a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \dots + a_n(x)y = 0.$$

If $a_0(x) \neq 0$ on some interval, then dividing all the terms of the equation by the coefficient $a_0(x)$ gives

$$y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = 0, \tag{18.11}$$

or

$$y^{(n)} = -p_1(x)y^{(n-1)} - \dots - p_n(x)y. \tag{18.12}$$

If $p_k(x)$, $k = 1, 2, \dots, n$, in (18.11) are continuous on the interval $[a, b]$, then the right-hand side of (18.12) is continuous in x, $a \leqslant x \leqslant b$, and in $y, y', \dots, y^{(n-1)}$ for any values of $y, y', \dots, y^{(n-1)}$, and it also has partial derivatives with respect to $y^{(k)}$ equal to $-p_{n-k}(x)$ and bounded on $[a, b]$. Therefore, by Theorem 18.1, we will deduce that if the coefficients $p_k(x)$, $k = 1, 2, \dots, n$, of (18.11) are continuous on $[a, b]$, then for any initial con-

ditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \ \ldots, \ y^{(n-1)}|_{x=x_0} = y_0^{(n-1)},$$
$$x_0 \in (a, b), \quad -\infty < y_0^{(k)} < +\infty, \quad k = 0, 1, \ldots, n-1$$

there exists a unique solution of (18.11) satisfying these initial conditions.

Recall the following notion. We say that on a set E an operator A with values in a set $\mathcal{F}$ is specified, if corresponding to each element $y \in E$ by a certain law is an element $f = Ay \in \mathcal{F}$. The set E is called the *range* of A.

Let E be a linear space. An operator A defined on E is said to be *linear*, if it is additive and homogeneous, i.e.,

(1) $A(y_1 + y_2) = Ay_1 + Ay_2 \quad \forall y_1, \forall y_2 \in E$;

(2) $A(\alpha y) = \alpha Ay \ \forall y \in E, \forall \alpha$ where α is a number.

We now represent (18.11) in the form

$$L[y] = 0,$$

where $L[y] \equiv y^{(n)} + p_1(x)y^{(n-1)} + \ldots + p_n(x)y$. It is easily seen that L is a linear differential operator defined at any rate in the linear space of the functions $y(x)$, which are continuous on (a, b), together with all the derivatives through the nth order.

The differential nature of the operator is obvious.

We will now show that it is linear, i.e., that

(1) $L[y_1 + y_2] = L[y_1] + L[y_2]$,

(2) $L[Cy] = CL[y]$,

where C is a constant. Really,

$$\begin{aligned} L[y_1 + y_2] &= (y_1 + y_2)^{(n)} + p_1(x)(y_1 + y_2)^{(n-1)} \\ &+ \ldots + p_n(x)(y_1 + y_2) = (y_1^{(n)} + p_1(x)y_1^{(n-1)} + \ldots + p_n(x)y_1) \\ &+ (y_2^{(n)} + p_1(x)y_2^{(n-1)} + \ldots + p_n(x)y_2) = L[y_1] + L[y_2]. \end{aligned}$$

Further,

$$\begin{aligned} L[Cy] &= Cy^{(n)} + p_1(x)Cy^{(n-1)} + \ldots + p_n(x)\,Cy \\ &= C(y^{(n)} + p_1(x)y^{(n-1)} + \ldots + p_n(x)y) = CL[y]. \end{aligned}$$

This suggests that

$$L\left[\sum_{i=1}^{m} C_i y_i(x)\right] = \sum_{i=1}^{m} C_i L[y_i],$$

where C_i are constants.

We will now establish some properties of the solutions of a linear homogeneous equation.

Theorem 18.2. *If a function $y_0(x)$ is a solution of a linear homogeneous differential equation $L[y] = 0$, then the function $Cy_0(x)$, where C is an arbitrary constant, is also a solution of the equation.*

◄ As stated, $L[y_0] \equiv 0$. We have to prove that $L[Cy_0] \equiv 0$. Using the homogeneity property of $L[y]$, we have $L[Cy_0] = CL[y_0] \equiv 0$. This means that $Cy_0(x)$ is a solution of the equation $L[y] = 0$. ►

Theorem 18.3. *If functions $y_1(x)$ and $y_2(x)$ are solutions of a linear homogeneous equation $L[y] = 0$, then the sum $y_1(x) + y_2(x)$ is also a solution of the equation.*

◄ As stated, $L[y_1] \equiv 0$ and $L[y_2] \equiv 0$. We prove that $L[y_1 + y_2] \equiv 0$. This immediately follows from the additivity property of $L[y]$:

$$L[y_1 + y_2] = L[y_1] + L[y_2] \equiv 0. \text{ ►}$$

Corollary. A linear combination with arbitrary constant coefficients $\sum_{i=1}^{m} C_i y_i(x)$ of solutions $y_1(x)$, $y_2(x)$, ..., $y_m(x)$ of the linear homogeneous differential equation $L[y] = 0$ is a solution of the equation.

The equation $L[y] = 0$ always has the trivial solution $y \equiv 0$. From Theorems 18.2 and 18.3 we obtain that a collection of solutions of $L[y] = 0$ forms a linear space, whose zero is the function $y \equiv 0$.

Theorem 18.4. *If a linear homogeneous equation $L[y] = 0$ with real coefficients $p_k(x)$, $k = 1, 2, ..., n$, has a complex solution $y(x) = u(x) + iv(x)$, then the real part of the solution $u(x)$ and its imaginary part $v(x)$ are each a solution of the equation.*

◄ Let $L[u + iv] \equiv 0$. We prove that $L[u] \equiv 0$ and $L[v] \equiv 0$. Using the linearity property of L, we get

$$L[u + iv] = L[u] + iL[v] \equiv 0.$$

It follows that $L[u] \equiv 0$ and $L[v] \equiv 0$, since the complex-valued function of a real argument becomes identically zero if and only if its real and imaginary parts are identically zero. ►

18.4 Linearly Dependent and Linearly Independent Systems of Functions

Consider a system of functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ defined on an interval (a, b).

Definition. We will say that the functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ are *linearly dependent* on the interval $a < x < b$, if there exist constants $\alpha_1, \alpha_2, ..., \alpha_n$ such that on that interval we have

$$\alpha_1 y_1(x) + \alpha_2 y_2(x) + ... + \alpha_n y_n(x) \equiv 0,$$

where at least one of α_i is nonzero.

If the identity holds only at $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, then the functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ are said to be *linearly independent* on the interval (a, b).

Consider now some examples of linearly dependent and linearly independent functions.

(1) The functions $y_1(x) = x$ and $y_2(x) = 2x$ are linearly dependent on any interval (a, b), since we have, for example, the identity $2y_1 - y_2 = 2x - 2x \equiv 0$, where $\alpha_1 = 2$, $\alpha_2 = -1$.

(2) The functions $1, x, x^2, \ldots, x^n$ are linearly independent on any interval (a, b), since the identity

$$\alpha_0 \cdot 1 + \alpha_1 x + \ldots + \alpha_n x^n \equiv 0, \quad x \in (a, b)$$

is possible only if $\alpha_i = 0$, $i = 0, 1, 2, \ldots, n$.

If at least one of α_i were nonzero, then in the left-hand side of the identity we would have a polynomial of degree not higher than n, which would have not more than n various roots and, consequently, would vanish at no more than n points in the interval.

(3) The functions $e^{k_1 x}, e^{k_2 x}, \ldots, e^{k_n x}$, where $k_i \neq k_j$ at $i \neq j$, are linearly independent on any interval (a, b). For simplicity we will confine ourselves to the case of $n = 3$. Suppose that the functions $e^{k_1 x}, e^{k_2 x}, e^{k_3 x}$ are linearly dependent. Then

$$\alpha_1 e^{k_1 x} + \alpha_2 e^{k_2 x} + \alpha_3 e^{k_3 x} \equiv 0.$$

Here at least one of α_i is nonzero. Suppose for definiteness that $\alpha_3 \neq 0$. If we divide the identity by $e^{k_1 x}$ and differentiate the resultant expression, we will get

$$\alpha_2(k_2 - k_1)\, e^{(k_2 - k_1)x} + \alpha_3(k_3 - k_1)\, e^{(k_3 - k_1)x} \equiv 0.$$

If then we divide this by $e^{(k_2 - k_1)x}$ and again differentiate with respect to x, we will find

$$\alpha_3(k_3 - k_1)(k_3 - k_2)\, e^{(k_3 - k_2)x} \equiv 0,$$

which is impossible, since $\alpha_3 \neq 0$ according to the assumption and $k_i \neq k_j$ for $i \neq j$. It follows that our assumption is wrong and the functions under consideration are linearly independent.

Remark. Linear dependence of a pair of functions implies that one of the functions is obtained from the other by multiplying by a constant, i.e.,

$$y_2(x) = \beta y_1(x), \quad \beta = \text{const.}$$

In general, if $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ are linearly dependent on (a, b), then at least one of them is a linear combination of the others.

Problem. Show that if a system of functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ is linearly independent on the interval (a, b), then any subsystem of this system of functions is also linearly independent on (a, b).

Theorem 18.5 (necessary condition for linear dependence of functions). *If functions $y_1(x), y_2(x), \ldots, y_n(x)$ that have derivatives through the order $n-1$ are linearly dependent on an interval (a, b), then on the interval the determinant*

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) & & y_n(x) \\ y_1'(x) & y_2'(x) & \ldots & y_n'(x) \\ \cdots & \cdots & \cdots & \cdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \ldots & y_n^{(n-1)}(x) \end{vmatrix}$$

called the Wronskian determinant (or simply *Wronskian*) *of the system of functions $y_1(x), y_2(x), \ldots, y_n(x)$ is identically zero, i.e.,*

$$W(x) \equiv 0 \quad in \quad (a, b).$$

◄ Suppose that $n = 3$ and the functions $y_1(x)$, $y_2(x)$ and $y_3(x)$ that are twice differentiable are linearly dependent on the interval (a, b). This means that in (a, b) we have

$$\alpha_1 y_1(x) + \alpha_2 y_2(x) + \alpha_3 y_3(x) \equiv 0,$$

where not all α_i $(i = 1, 2, 3)$ are zero. To be more specific, let $\alpha_1 \neq 0$. We solve the identity for $y_1(x)$ and differentiate it twice

$$y_1(x) = -\frac{\alpha_2}{\alpha_1} y_2(x) - \frac{\alpha_3}{\alpha_1} y_3(x), \tag{18.13}$$

$$\begin{aligned} y_1'(x) &= -\frac{\alpha_2}{\alpha_1} y_2'(x) - \frac{\alpha_3}{\alpha_1} y_3'(x), \\ y_1''(x) &= -\frac{\alpha_2}{\alpha_1} y_2''(x) - \frac{\alpha_3}{\alpha_1} y_3''(x). \end{aligned} \tag{18.14}$$

Let us form the Wronskian of the system $y_1(x)$, $y_2(x)$ and $y_3(x)$

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) & y_3(x) \\ y_1'(x) & y_2'(x) & y_3'(x) \\ y_1''(x) & y_2''(x) & y_3''(x) \end{vmatrix},$$

or, by (18.13) and (18.14),

$$W(x) = \begin{vmatrix} -\frac{\alpha_2}{\alpha_1} y_2(x) - \frac{\alpha_3}{\alpha_1} y_3(x) & y_2(x) & y_3(x) \\ -\frac{\alpha_2}{\alpha_1} y_2'(x) - \frac{\alpha_3}{\alpha_1} y_3'(x) & y_2'(x) & y_3'(x) \\ -\frac{\alpha_2}{\alpha_1} y_2''(x) - \frac{\alpha_3}{\alpha_1} y_3''(x) & y_2''(x) & y_3''(x) \end{vmatrix}.$$

The first column of the determinant is a linear combination of the other two for any $x \in (a, b)$. Such a determinant is known to be zero. Hence $W(x) = 0\ \forall x \in (a, b)$. ►

We easily prove the following theorem by contradiction.

Theorem 18.6. *If the Wronskian $W(x)$ of a system of n functions is not identically equal to zero in a certain interval (a, b), then these functions are linearly independent on the interval.*

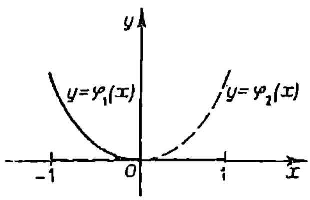


Fig. 18.2

For an arbitrary system of functions that can be differentiated $n - 1$ times on (a, b) the reverse of Theorem 18.5 is not true. To see this, consider an example. For the functions (Fig. 18.2)

$$\varphi_1(x) = \begin{cases} x^2, & -1 < x < 0, \\ 0, & 0 \leqslant x < 1, \end{cases} \qquad \varphi_2(x) = \begin{cases} 0, & -1 < x < 0, \\ x^2, & 0 \leqslant x < 1, \end{cases}$$

the Wronskian in the interval $(-1, 1)$ is identically zero:

$$W(x) = \begin{vmatrix} \varphi_1 & \varphi_2 \\ \varphi_1' & \varphi_2' \end{vmatrix} = \begin{cases} \begin{vmatrix} x^2 & 0 \\ 2x & 0 \end{vmatrix} \equiv 0, & -1 < x < 0, \\ \begin{vmatrix} 0 & x^2 \\ 0 & 2x \end{vmatrix} \equiv 0, & 0 \leqslant x < 1. \end{cases}$$

As is easily seen, however, the functions $\varphi_1(x)$ and $\varphi_2(x)$ are linearly independent on the interval $(-1, 1)$. Notice that on the intervals $(-1, 0)$ and $(0, 1)$ the functions $\varphi_1(x)$ and $\varphi_2(x)$ are linearly dependent.

We generalize somewhat the above example:

$$\psi_1(x) = \begin{cases} x^m, & x < 0, \\ 0, & x \geqslant 0, \end{cases} \qquad \psi_2(x) = \begin{cases} 0, & x < 0, \\ x^m, & x \geqslant 0 \end{cases} \quad (m > 1, \ \text{integer}).$$

These functions are linearly independent on any interval that contains the point $x = 0$, at the same time their Wronskian is identically equal to zero. Also, the function $\psi_2(x)$ has everywhere continuous derivatives, up to the order $m - 1$, and only the derivative of the mth order has a discontinuity with a finite jump at $x = 0$. Taking m sufficiently large, we obtain a system of functions having continuous derivatives of any desired order.

Problem. What can be said about the Wronskian of a system of functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$, if we only know that these functions are (a) linearly dependent or (b) linearly independent?

Theorem 18.7 (necessary condition for linear independence of solutions). *If functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ are linearly independent on the interval (a, b) and are solutions of the linear homogeneous differential equation*

$$y^{(n)} + p_1(x)y^{(n-1)} + \ldots + p_n(x)y = 0 \tag{18.15}$$

with coefficients $p_k(x)$ continuous on $[a, b]$, then the Wronskian of the system of solutions

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) & & y_n(x) \\ y_1'(x) & y_2'(x) & & y_n'(x) \\ \dots & \dots & \dots & \dots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \ldots & y_n^{(n-1)}(x) \end{vmatrix}$$

cannot vanish at any point in (a, b).

◄ Let $n = 3$. Suppose that at a point $x_0 \in (a, b)$ the Wronskian is zero, i.e., $W(x_0) = 0$. We set up a system of three linear homogeneous algebraic equations in α_1, α_2, and α_3:

$$\begin{cases} \alpha_1 y_1(x_0) + \alpha_2 y_2(x_0) + \alpha_3 y_3(x_0) = 0, \\ \alpha_1 y_1'(x_0) + \alpha_2 y_2'(x_0) + \alpha_3 y_3'(x_0) = 0, \\ \alpha_1 y_1''(x_0) + \alpha_2 y_2''(x_0) + \alpha_3 y_3''(x_0) = 0. \end{cases} \tag{18.16}$$

For this system, $W(x_0)$ is assumed to be zero; therefore, the system has a nonzero solution $\tilde{\alpha}_1$, $\tilde{\alpha}_2$, and $\tilde{\alpha}_3$, i.e., at least one of $\tilde{\alpha}_i$ is nonzero.

Consider the function

$$y = \tilde{\alpha}_1 y_1(x) + \tilde{\alpha}_2 y_2(x) + \tilde{\alpha}_3 y_3(x). \tag{18.17}$$

It is a linear combination of the solutions $y_1(x)$, $y_2(x)$, and $y_3(x)$ of (18.15), and hence it is itself a solution of the equation.

By (18.16), this solution satisfies the zero initial conditions

$$\begin{aligned} y(x_0) &= \tilde{\alpha}_1 y_1(x_0) + \tilde{\alpha}_2 y_2(x_0) + \tilde{\alpha}_3 y_3(x_0) = 0, \\ y'(x_0) &= \tilde{\alpha}_1 y_1'(x_0) + \tilde{\alpha}_2 y_2'(x_0) + \tilde{\alpha}_3 y_3'(x_0) = 0, \\ y''(x_0) &= \tilde{\alpha}_1 y_1''(x_0) + \tilde{\alpha}_2 y_2''(x_0) + \tilde{\alpha}_3 y_3''(x_0) = 0. \end{aligned} \tag{18.18}$$

Such initial conditions are obviously met by the trivial solution $y \equiv 0$ of (18.15) and, by virtue of the theorem on uniqueness of solution, by this solution alone. Accordingly, $\tilde{\alpha}_1 y_1(x) + \tilde{\alpha}_2 y_2(x) + \tilde{\alpha}_3 y_3(x) \equiv 0$ on (a, b) at least one of $\tilde{\alpha}_i$ being nonzero. It follows that the solutions $y_1(x)$, $y_2(x)$, and $y_3(x)$ appear to be linearly dependent, despite the conditions of the

theorem. The contradiction emerged due to the assumption that $W(x)$ turns into zero at $x_0 \in (a, b)$. And so our assumption is wrong and $W(x) \neq 0$ everywhere in (a, b).

From Theorems 18.5 and 18.7 we will deduce, as consequence, the following important theorem.

Theorem 18.8. *Let $y_1(x), y_2(x), \ldots, y_n(x)$ be particular solutions of* (18.15) *with coefficients continuous in the interval (a, b). For them to be linearly independent in (a, b) it is necessary and sufficient that the Wronskian $W(x)$ of the system of solutions be nonzero.*

◄ The necessity follows immediately from Theorem 18.7.

The sufficiency follows from the fact that if the functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ are linearly dependent, then by Theorem 18.5 we have $W(x) \equiv 0$. Therefore, if $W(x) \neq 0$, then the functions $y_1(x), y_2(x), \ldots, y_n(x)$ cannot be linearly dependent, i.e., they are linearly independent. ►

Problems. (1) Given the equation $y'' + p_1(x)y' + p_2(x)y = 0$ with continuous coefficients, prove that two solutions of the equation that have a maximum at the same value of x are linearly dependent.

(2) Given the equation $y'' + p_1(x)y' + p_2(x)y = 0$ with continuous coefficients, prove that the ratio of any two linearly independent solutions of the equation cannot have any points of maximum.

(3) Given the equation $y'' + p_1(x)y' + p_2(x)y = 0$ with coefficients continuous on the interval $[a, b]$, show that two linearly independent solutions $y_1(x)$ and $y_2(x)$ of the equation cannot turn into zero at the same value of $x_0 \in (a, b)$.

18.5 Structure of the General Solution of a Linear Homogeneous Differential Equation

Theorem 18.9 (on the structure of the general solution of a linear homogeneous differential equation). *A linear homogeneous differential equation*

$$y^{(n)} + p_1(x)y^{(n-1)} + \ldots + p_n(x)y = 0 \tag{18.19}$$

with coefficients $p_k(x)$, $k = 1, 2, \ldots, n$, continuous on the interval $[a, b]$ has in the domain $a < x < b$, $|y^{(k)}| < +\infty$, $k = 0, 1, 2, n - 1$, the general solution

$$y = \sum_{i=1}^{n} C_i y_i(x), \tag{18.20}$$

i.e., a linear combination of n particular solutions $y_i(x)$, $i = 1, 2, \ldots, n$, which are linearly independent in (a, b). Here $C_1, C_2, \ldots, C_n$ are arbitrary constants.

◄ We will base ourselves on the definition of the general solution and will simply check whether the family of functions $y = \sum_{i=1}^{n} C_i y_i(x)$ satisfies the conditions (1) and (2) of that definition.

For any $C_1, C_2, \ldots, C_n$ the function $y(x)$, given by (18.20), is a solution of (18.19). This follows from the fact that, as it has been established above, any linear combination of particular solutions of a linear homogeneous equation is itself another solution of the equation.

For (18.19) the conditions of Theorem 18.1 are met for $x \in [a, b]$, therefore, we have only to show that we can always choose $C_1, C_2, \ldots, C_n$ to satisfy the arbitrary initial conditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \ \ldots, \ y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}$$

where $x_0 \in (a, b)$.

We will only consider the case of $n = 3$. We require that the solution

$$y(x) = C_1 y_1(x) + C_2 y_2(x) + C_3 y_3(x)$$

satisfy the initial conditions. We thus obtain the system of three linear algebraic equations in C_1, C_2, C_3

$$\begin{cases} C_1 y_1(x_0) + C_2 y_2(x_0) + C_3 y_3(x_0) = y_0, \\ C_1 y_1'(x_0) + C_2 y_2'(x_0) + C_3 y_3'(x_0) = y_0', \\ C_1 y_1''(x_0) + C_2 y_2''(x_0) + C_3 y_3''(x_0) = y_0''. \end{cases} \tag{18.21}$$

The determinant of the system is the Wronskian $W(x_0)$ of the linearly independent system of solutions of (18.19), and hence it is nonzero at any $x \in (a, b)$, specifically at $x = x_0$. Therefore, the system (18.21) can be solved uniquely for C_1, C_2, C_3 at any $x_0 \in (a, b)$ and any right-hand sides, i.e., at any y_0, y_0', y_0''. And this means that it is possible to choose C_1^0, C_2^0, C_3^0 such that the particular solution $y(x) = C_1^0 y_1(x) + C_2^0 y_2(x) + C_3^0 y_3(x)$ would meet the initial conditions whatever they might be. ►

It follows from Theorem 18.9 that if n linearly independent particular solutions of a linear homogeneous differential equation of nth order are known, then any other solution of the equation can be represented as a linear combination of these particular solutions, and hence it will be linearly dependent with respect to them. And so the maximum number of linearly independent solutions of a homogeneous linear differential equation is equal to its order.

In summary, *the collection of solutions of a linear homogeneous differential equation forms a linear space whose dimension is equal to the order of the equation.*

We introduce the notion of the fundamental set of solutions.

Definition. A collection of any n linearly independent particular solutions of a linear homogeneous differential equation of nth order is called its *fundamental set of solutions.*

Theorem 18.10. *For each linear homogeneous equation* (18.19) *with continuous coefficients* $p_k(x)$ *there exists a fundamental set of solutions* (*and even an infinite number of fundamental sets of solutions*).

◄ Consider, for example, the homogeneous equation of the second order

$$y'' + p_1(x)y' + p_2(x)y = 0 \tag{18.22}$$

with coefficients continuous on the interval $[a, b]$. Let $x_0 \in (a, b)$. By Theorem 18.1 equation (18.22) has the solutions

$$y = y_1(x), \qquad y = y_2(x) \tag{18.23}$$

that at $x = x_0$ satisfy the initial conditions

$$y_1(x_0) = 1, \quad y_1'(x_0) = 0, \quad y_2(x_0) = 0, \quad y_2'(x_0) = 1. \tag{18.23'}$$

The Wronskian of (18.23) at the point x_0 is nonzero:

$$W(x_0) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \neq 0.$$

Consequently, the set of solutions (18.23) of equation (18.22) is fundamental. The choice of initial conditions (18.23′) led to the formation of one fundamental set. For initial data at point x_0 we can take any set of numbers $y_1(x_0) = a_{11}, y_1'(x_0) = a_{21}, y_2(x_0) = a_{12}, y_2'(x_0) = a_{22}$, provided that the Wronskian

$$W(x_0) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

is different from zero. Clearly, we can find an infinite variety of such sets and so we can construct an infinite number of fundamental sets of solutions for (18.22).

Problem. Set up the general solution for the equation $y' + p(x)y = 0$, given the nonzero particular solution $y_1(x)$ of the equation.

Theorem 18.11. *If two equations of the form*

$$y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = 0$$

and

$$y^{(n)} + q_1(x)y^{(n-1)} + \dots + q_n(x)y = 0,$$

where the functions $p_i(x)$ *and* $q_i(x)$, $i = 1, 2, \dots, n$, *are continuous on the interval* $[a, b]$, *have a common fundamental set of solutions* $y_1(x), y_2(x), \dots, y_n(x)$, *then these equations coincide, i.e.,* $p_i(x) \equiv q_i(x)$ *on* $[a, b]$.

A fundamental set of solutions thus completely defines the linear homogeneous equation (18.19), i.e., completely defines the coefficients $p_i(x)$, $i = 1, 2, ..., n$, of the equation.

Correspondingly, we can formulate the problem of finding an equation of the form (18.19) that has a given fundamental set of solutions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$. We now represent the left-hand side of the equation in the form of a determinant:

$$\begin{vmatrix} y_1(x) & y_2(x) & \dots & y_n(x) & y(x) \\ y_1'(x) & y_2'(x) & \dots & y_n'(x) & y'(x) \\ \dots & \dots & \dots & \dots & \dots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \dots & y_n^{(n-1)}(x) & y^{(n-1)}(x) \\ y_1^{(n)}(x) & y_2^{(n)}(x) & \dots & y_n^{(n)}(x) & y^{(n)}(x) \end{vmatrix} = 0, \qquad (18.24)$$

where $y(x)$ is the desired function, and $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ are a given fundamental set of solutions.

Equation (18.24) has solutions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$, since if we substitute any of these n functions for $y(x)$ two columns of the determinant become identically equal and the determinant will become identically zero in $x \in (a, b)$. Expanding the determinant with respect to the elements of the last column we obtain from (18.24) an equation of the form

$$W(x)\,y^{(n)} - W_1(x)\,y^{(n-1)} + \dots \pm W_n(x)\,y = 0, \qquad (18.25)$$

where $W(x)$ is the Wronskian of the set of functions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$, and

$$W_1(x) = \begin{vmatrix} y_1(x) & y_2(x) & \dots & y_n(x) \\ y_1'(x) & y_2'(x) & \dots & y_n'(x) \\ \dots & \dots & \dots & \dots \\ y_1^{(n-2)}(x) & y_2^{(n-2)}(x) & \dots & y_n^{(n-2)}(x) \\ y_1^{(n)}(x) & y_2^{(n)}(x) & \dots & y_n^{(n)}(x) \end{vmatrix}.$$

The Wronskian $W(x)$ of the fundamental set of solutions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ is distinct from zero in the entire interval (a, b). Dividing all the terms of (18.25) by $W(x) \neq 0$, we will reduce the equation to the form (18.19)

$$y^{(n)} + p_1(x)\,y^{(n-1)} + \dots + p_n(x)\,y = 0,$$

where, in particular, $p_1(x) = -W_1(x)/W(x)$.

It can be shown that if the elements a_{ij} of the determinant Δ of nth order are differentiable functions of x, i.e., $a_{ij} = a_{ij}(x)$, then the derivative of the determinant $\Delta = \Delta(x)$ is the sum of n determinants

$$\Delta' = \Delta_1' + \Delta_2' + \dots + \Delta_n',$$

where Δ'_k, $k = 1, 2, \ldots, n$, is the determinant derived from the original one by replacing the elements of its kth row by the derivatives of respective elements. For example, for the Wronskian of the set $y_1(x)$ and $y_2(x)$ we have

$$\frac{d}{dx}\begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} = \begin{vmatrix} y_1'(x) & y_2'(x) \\ y_1'(x) & y_2'(x) \end{vmatrix}$$
$$+ \begin{vmatrix} y_1(x) & y_2(x) \\ y_1''(x) & y_2''(x) \end{vmatrix} = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1''(x) & y_2''(x) \end{vmatrix}.$$

It can be readily verified that $W_1(x) = dW(x)/dx$, and so

$$p_1(x) = -\frac{W'(x)}{W(x)}.$$

Integrating the last identity in x from x_0 to x we will arrive at the Ostrogradsky-Liouville formula

$$W(x) = W(x_0)\,e^{-\int_{x_0}^{x} p_1(x)\,dx}. \tag{18.26}$$

Problem. Form the linear differential equation of the second order whose solutions are $y_1 = x$ and $y_2 = x^2$. Show that the functions x and x^2 are linearly independent on the interval $(-\infty, +\infty)$. Make sure that the Wronskian for these functions is zero at $x = 0$. Why is this not at variance with the necessary condition for linear independence of a set of solutions of a linear homogeneous differential equation?

18.6 Linear Homogeneous Differential Equations with Constant Coefficients

Let

$$y'' + p_1y' + p_2y = 0, \tag{18.27}$$

where p_1 and p_2 are real numbers, be a linear homogeneous differential equation of the *second order*. To find the general solution of the equation we need to find two of its linearly independent particular solutions.

According to Euler, we will seek them in the form

$$y = e^{\lambda x}, \tag{18.28}$$

where $\lambda = \text{const}$. Then $y' = \lambda e^{\lambda x}$, $y' = \lambda^2 e^{\lambda x}$. Substituting these expressions for y and its derivatives into (18.27), we get

$$e^{\lambda x}(\lambda^2 + p_1\lambda + p_2) = 0.$$

Since $e^{\lambda x} \neq 0$, we must have $\lambda^2 + p_1\lambda + p_2 = 0$. Consequently, the function $y = e^{\lambda x}$ will be a solution to (18.27), i.e., it will turn it into identity in x, if λ obeys the algebraic equation

$$\lambda^2 + p_1\lambda + p_2 = 0. \tag{18.29}$$

Equation (18.29) is called the *characteristic equation* in relation to (18.27), and its left-hand side $\varphi(\lambda) \equiv \lambda^2 + p_1\lambda + p_2$ is called the *characteristic polynomial*.

Equation (18.29) is a quadratic equation. We denote its roots by λ_1 and λ_2. They may be (1) real and different; (2) complex; (3) real and equal.

Consider each case separately.

(1) If the roots λ_1 and λ_2 of the characteristic equation are real and different, then the particular solutions of (18.27) will be

$$y_1 = e^{\lambda_1 x}, \quad y_2 = e^{\lambda_2 x}.$$

These solutions are linearly independent ($\lambda_1 \neq \lambda_2$), and so they form a fundamental set of solutions of the equation. The general solution of the equation has the form

$$y = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x},$$

where C_1 and C_2 are arbitrary constants.

Example. Find the general solution of the equation

$$y'' - 3y' + 2y = 0.$$

◄ We set up the characteristic equation

$$\lambda^2 - 3\lambda + 2 = 0.$$

It has the roots $\lambda_1 = 1$ and $\lambda_2 = 2$. The desired general solution will be

$$y = C_1 e^{x} + C_2 e^{2x}. ►$$

(2) Let the roots of the characteristic equation be complex. Since the coefficients p_1 and p_2 of the characteristic equation are real, the complex roots are pairwise conjugate. Suppose that $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$. Particular solutions of (18.27) can be written in the form

$$\tilde{y}_1 = e^{(\alpha + i\beta)x}, \quad \tilde{y}_2 = e^{(\alpha - i\beta)x}.$$

These are complex-valued functions of a real argument x, and we will only be concerned with real solutions. Using the Euler formulas

$$e^{i\beta x} = \cos\beta x + i\sin\beta x, \quad e^{-i\beta x} = \cos\beta x - i\sin\beta x,$$

we can represent the particular solutions $\tilde{y}_1$ and $\tilde{y}_2$ of (18.27) in the form

$$\tilde{y}_1 = e^{\alpha x}(\cos\beta x + i\sin\beta x), \quad \tilde{y}_2 = e^{\alpha x}(\cos\beta x - i\sin\beta x).$$

Using Theorem 18.4 we find that other particular solutions of (18.27) will be

$$y_1 = e^{\alpha x}\cos\beta x, \quad y_2 = e^{\alpha x}\sin\beta x.$$

These solutions are linearly independent, since $y_2(x)/y_1(x) = \tan\beta x \not\equiv \text{const}$, and so they form a fundamental set of solutions. The general solution of (18.27) in this case has the form

$$y = C_1 e^{\alpha x}\cos\beta x + C_2 e^{\alpha x}\sin\beta x.$$

Example. Find the general solution of the equation

$$y'' + 2y' + 5y = 0.$$

◄ We write the characteristic equation

$$\lambda^2 + 2\lambda + 5 = 0$$

Its roots are $\lambda_1 = -1 + 2i$, and $\lambda_2 = -1 - 2i$, therefore $\alpha = -1$, $\beta = 2$. The general solution will be

$$y = C_1 e^{-x}\cos 2x + C_2 e^{-x}\sin 2x. \text{ ►}$$

(3) Suppose now that the roots of the characteristic equation are real and equal. One particular solution $y_1 = e^{\lambda_1 x}$ follows immediately. The other particular solution, which is linearly independent of the first one, will be sought in the form

$$y_2 = e^{\lambda_1 x} u(x),$$

where $u(x)$ is a new unknown function. Differentiating gives

$$y_2' = \lambda_1 e^{\lambda_1 x} u + e^{\lambda_1 x} u', \quad y_2'' = \lambda_1^2 e^{\lambda_1 x} u + 2\lambda_1 e^{\lambda_1 x} u' + e^{\lambda_1 x} u''.$$

Substituting the expressions obtained into (18.27) gives

$$e^{\lambda_1 x}[u'' + (2\lambda_1 + p_1)u' + (\lambda_1^2 + p_1\lambda_1 + p_2)u] = 0. \tag{18.30}$$

Since λ_1 is a root of the characteristic equation,

$$\lambda_1^2 + p_1\lambda_1 + p_2 = 0,$$

and since λ_1 is a double root, we also have

$$2\lambda_1 + p_1 = 0.$$

Relation (18.30) thus becomes $u'' = 0$. Hence $u = Ax + B$, where A and B are constants. We can, in particular, put $A = 1$ and $B = 0$, then $u = x$.

For the second particular solution of the equation we can thus take $y_2 = xe^{\lambda_1 x}$. This solution is linearly independent of the first one, since $y_2(x)/y_1(x) = x \not\equiv \text{const}$. The solutions $y_1 = e^{\lambda_1 x}$ and $y_2 = xe^{\lambda_1 x}$ form a fundamental set of solutions of (18.27) and the general solution will then be

$$y = C_1 e^{\lambda_1 x} + C_2 x e^{\lambda_1 x} \quad \text{or} \quad y = e^{\lambda_1 x}(C_1 + C_2 x).$$

Example. Find the general solution of the equation

$$y'' + 2y' + y = 0.$$

◄ The characteristic equation $\lambda^2 + 2\lambda + 1 = 0$ has multiple roots $\lambda_1 = \lambda_2 = -1$. Therefore, the general solution of the original differential equation will be

$$y = C_1 e^{-x} + C_2 x e^{-x}.$$ ►

Remark. Consider the linear homogeneous differential equation (in general, with variable coefficients)

$$L[y] = y^{(n)} + p_1(x)\, y^{(n-1)} + \ldots + p_n(x)\, y = 0. \tag{18.31}$$

Let $y_1(x)$ be a particular solution of the equation. We introduce a new desired function $u(x)$ by

$$y = y_1(x)\, u(x).$$

This equation is solvable for $u(x)$ in the intervals where $y_1(x)$ does not vanish. From this we find the derivatives with respect to y

$$y' = y_1 u' + y_1' u,$$

$$y'' = y_1 u'' + 2y_1' u' + y_1'' u,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$y^{(n)} = y_1 u^{(n)} + \frac{n}{1!} y_1' y^{(n-1)} + \ldots + y_1^{(n)} u$$

and substitute them into (18.31):

$$y_1(x)\, u^{(n)} + [ny_1' + p_1 y_1]\, u^{(n-1)} + \ldots + [\ldots]\, u'$$
$$+ [y_1^{(n)} + p_1 y_1^{(n-1)} + \ldots + p_n(x)\, y_1]\, u = 0.$$

For $u(x)$ we again obtain an equation of order n, but the coefficient at $u(x)$ is $L[y_1]$. It is identically equal to zero since $y_1(x)$ is a solution of (18.31). Accordingly, the order of the equation obtained will decrease if we introduce a new function $z(x) = u'(x)$. If we also divide all the terms of the last equation by $y_1(x) \neq 0$, we will reduce it to the form

$$z^{(n-1)} + q_1(x)\, z^{(n-2)} + \ldots + q_{n-1}(x)\, z = 0.$$

To summarize, if we know a particular solution of (18.31), the problem of integrating this equation reduces to integrating a linear homogeneous equation of order $n - 1$. It can be shown that if we know two particular linearly independent solutions, then the order of the equation can be reduced by two. In general, if we know r particular linearly independent solutions of a linear homogeneous differential equation, then the order of this equation can be reduced by r. ►

Linear differential equations with constant coefficients occur in problems concerned with mechanical and electrical oscillations. Consider the equation of free mechanical oscillations, where the independent variable is the time t:

$$m\frac{d^2y}{dt^2} + h\frac{dy}{dt} + ky = 0. \tag{18.32}$$

Here y is the displacement of the vibrating point about the position of equilibrium, m is the point mass, and h is the friction coefficient (we consider the friction force to be proportional to speed), $k > 0$ is the elasticity coefficient of the restoring force, which is believed to be proportional to the displacement. The characteristic equation for (18.32)

$$m\lambda^2 + h\lambda + k = 0$$

has the roots $\lambda_{1,2} = -h/2m \pm \sqrt{h^2/(4m^2) - k/m}$. If friction is sufficiently large, i.e., $h^2 > 4mk$, then these roots are real and negative. The general solution of (18.32) will then be

$$y = C_1e^{\lambda_1 t} + C_2e^{\lambda_2 t}. \tag{18.33}$$

Since $\lambda_1 < 0$ and $\lambda_2 < 0$, then from (18.33) we conclude that when friction is large the displacement tends to zero with increasing t, without vibrations.

If friction is small, i.e., $h^2 < 4mk$, then the characteristic equation has complex conjugate roots $-\alpha \pm i\beta$, where $\alpha = h/(2m) > 0$, and $\beta = \sqrt{k/m - h^2/(4m^2)}$. The general solution of (18.33) will then be given by

$$y = C_1e^{-\alpha t}\cos\beta t + C_2e^{-\alpha t}\sin\beta t \quad (\alpha > 0)$$

or

$$y = Ae^{-\alpha t}\sin(\beta t + \delta) \quad (A,\ \delta = \text{const}).$$

It is seen from this that when friction is small the vibrations are damped.

Suppose now that we have no friction, i.e., that $h = 0$. In that case, the characteristic equation $m\lambda^2 + k = 0$ has purely imaginary roots $\lambda_{1,2} = \pm i\sqrt{k/m}$. The solution of (18.32) has the form

$$y = C_1\cos\omega t + C_2\sin\omega t = A\sin(\omega t + \delta),$$

where $\omega = \sqrt{k/m}$, i.e., we now have undamped harmonic oscillations with the frequency $\omega = \sqrt{k/m}$ and arbitrary amplitude A and initial phase δ.

Problem. Consider the equation

$$y'' + p_1y' + p_2y = 0,$$

where p_1 and p_2 are constants. At what p_1 and p_2 (1) all the solutions of the equation tend to zero as $x \to +\infty$ and (2) each solution of the equation becomes zero on an infinite set of points x?

We now consider the linear homogeneous differential equaton of *arbitrary order* $n \geqslant 1$ with constant coefficients

$$L[y] \equiv y^{(n)} + p_1 y^{(n-1)} + \ldots + p_n y = 0, \tag{18.34}$$

where $p_1, p_2, \ldots, p_n$ are real numbers. The general solution of (18.34) is sought in the same manner as for the second-order equation.

(1) We look for the solution in the form $y = e^{\lambda x}$. Substituting $e^{\lambda x}$ for y in (18.34) gives

$$L[e^{\lambda x}] = e^{\lambda x}\varphi(\lambda) = 0.$$

This leads to the characteristic equation

$$\varphi(\lambda) \equiv \lambda^n + p_1\lambda^{n-1} + \ldots + p_n = 0. \tag{18.35}$$

(2) We find the roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the characteristic equation.

(3) According to the nature of the roots we write the particular linearly independent solutions of (18.34) proceeding from the fact that

(a) corresponding to each real single root λ of the characteristic equation is a particular solution $e^{\lambda x}$ of (18.34);

(b) corresponding to each pair of single complex conjugate roots $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$ are two linearly independent particular solutions $e^{\alpha x}\cos\beta x$ and $e^{\alpha x}\sin\beta x$ of (18.34);

(c) corresponding to each real root λ of multiplicity r are r linearly independent particular solutions $e^{\lambda x}, xe^{\lambda x}, \ldots, x^{r-1}e^{\lambda x}$ of (18.34).

Indeed, let λ be a root of multiplicity r of the characteristic equation $\varphi(\lambda) = 0$. We will treat $y = e^{\lambda x}$ as a function of two arguments x and λ. It has continuous derivatives of all orders with respect to x and λ. More precisely, $\dfrac{\partial^m}{\partial\lambda^m}(e^{\lambda x}) = x^m e^{\lambda x}$. Therefore, partial derivatives of $y = e^{\lambda x}$ with respect to x and λ are independent of the order of differentiation (the operations of differentiating y with respect to x and λ are interchangeable), so that

$$L\left[\frac{\partial^m y}{\partial\lambda^m}\right] = \frac{\partial^m}{\partial\lambda^m} L[y].$$

Using this property, and also the fact that

$$L[e^{\lambda x}] = e^{\lambda x}\varphi(\lambda), \tag{18.36}$$

we obtain

$$\left.\begin{aligned}
L[xe^{\lambda x}] &= L\left[\frac{\partial}{\partial\lambda}e^{\lambda x}\right] = \frac{\partial}{\partial\lambda}L[e^{\lambda x}] = \frac{\partial}{\partial\lambda}(e^{\lambda x}\varphi(\lambda))\\
&= xe^{\lambda x}\varphi(\lambda) + e^{\lambda x}\varphi'(\lambda),\\
L[x^2e^{\lambda x}] &= L\left[\frac{\partial^2}{\partial\lambda^2}e^{\lambda x}\right] = \frac{\partial^2}{\partial\lambda^2}L[e^{\lambda x}] = \frac{\partial^2}{\partial\lambda^2}(e^{\lambda x}\varphi(\lambda))\\
&= x^2e^{\lambda x}\varphi(\lambda) + 2xe^{\lambda x}\varphi'(\lambda) + e^{\lambda x}\varphi''(\lambda),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
L[x^{r-1}e^{\lambda x}] &= \frac{\partial^{r-1}}{\partial\lambda^{r-1}}(e^{\lambda x}\varphi(\lambda)) = x^{r-1}e^{\lambda x}\varphi(\lambda)\\
&+ \frac{(r-1)}{1!}x^{r-2}e^{\lambda x}\varphi'(\lambda) + \ldots + e^{\lambda x}\varphi^{(r-1)}(\lambda).
\end{aligned}\right\} \quad (18.37)$$

If λ is an r-tuple root of the characteristic equation $\varphi(\lambda) = 0$, then $\varphi(\lambda) = \varphi'(\lambda) = \ldots = \varphi^{(r-1)}(\lambda) = 0$, $\varphi^{(r)}(\lambda) \neq 0$, and so the right-hand sides of (18.36) and (18.37) are identically equal to zero in x:

$$L[e^{\lambda x}] \equiv 0, \quad L[xe^{\lambda x}] \equiv 0, \quad \ldots, \quad L[x^{r-1}e^{\lambda x}] \equiv 0.$$

This means that $e^{\lambda x}$, $xe^{\lambda x}$, ..., $x^{r-1}e^{\lambda x}$ are solutions of (18.34). It can easily be verified that the functions $e^{\lambda x}$, $xe^{\lambda x}$, ..., $x^{r-1}e^{\lambda x}$ are linearly independent on any interval (a, b) where x is defined.

(d) what is said in (c) also holds true for complex roots. Therefore, corresponding to each pair of complex conjugate roots $\lambda = \alpha + i\beta$, and $\bar{\lambda} = \alpha - i\beta$ of multiplicity μ are 2μ particular solutions of (18.34)

$$\begin{aligned}
&e^{\alpha x}\cos\beta x, \quad xe^{\alpha x}\cos\beta x, \quad \ldots, \quad x^{\mu-1}e^{\alpha x}\cos\beta x,\\
&e^{\alpha x}\sin\beta x, \quad xe^{\alpha x}\sin\beta x, \quad \ldots, \quad x^{\mu-1}e^{\alpha x}\sin\beta x.
\end{aligned}$$

(4) The number of particular solutions of (18.34) constructed in such a manner is equal to the order n of the equation. It can be shown that these solutions are all linearly independent in combination. Given n linearly independent particular solutions $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ of equation (18.34), we obtain the general solution

$$y = C_1y_1(x) + C_2y_2(x) + \ldots + C_ny_n(x),$$

where C_1, C_2, ..., C_n are arbitrary constants.

Example. Find the general solution of the equation $y^{(6)} - y'' = 0$.

◄ (1) We write the characteristic equation

$$\lambda^6 - \lambda^2 = 0 \quad \text{or} \quad \lambda^2(\lambda^4 - 1) = 0.$$

(2) We find the roots of the characteristic equation

$$\lambda_1 = \lambda_2 = 0, \quad \lambda_3 = 1, \quad \lambda_4 = -1, \quad \lambda_5 = i, \quad \lambda_6 = -i.$$

(3) From the nature of the roots we write the particular linearly independent solutions of the equation:

$$y_1 = 1, \quad y_2 = x, \quad y_3 = e^x, \quad y_4 = e^{-x}, \quad y_5 = \cos x, \quad y_6 = \sin x.$$

(4) The general solution of the equation has the form

$$y = C_1 + C_2 x + C_3 e^x + C_4 e^{-x} + C_5 \cos x + C_6 \sin x. \ \blacktriangleright$$

Table 18.1

Differential equation

$$L[y] \equiv y^{(n)}(x) + p_1 y^{(n-1)}(x) + \ldots + p_n y(x) = 0$$

($p_1, p_2, \ldots, p_n$ are real numbers)

Characteristic equation

$$\varphi(\lambda) \equiv \lambda^n + p_1\lambda^{n-1} + \ldots + p_n = 0$$

Roots

$$\lambda_1, \lambda_2, \ldots, \lambda_n$$

of the characteristic equation $\varphi(\lambda) = 0$

Particular linearly independent solutions

$$y_1(x), y_2(x), \ldots, y_n(x)$$

of the differential equation $L[y] = 0$

General solution of the equation $L[y] = 0$

$$y(x) = C_1 y_1(x) + C_2 y_2(x) + \ldots + C_n y_n(x)$$

($C_1, C_2, \ldots, C_n$ are arbitrary constants)

18.7 Equations Reducible to Equations with Constant Coefficients

There are linear differential equations with variable coefficients that by a change of variables can be transformed into equations with constant coefficients. Among them is the Euler equation

$$x^n \frac{d^n y}{dx^n} + p_1 x^{n-1} \frac{d^{n-1} y}{dx^{n-1}} + \dots + p_{n-1} x \frac{dy}{dx} + p_n y = 0,$$

where $p_1, p_2, \dots, p_n$ are constants.

We will confine ourselves to the Euler equation of the second order (it occurs in problems of mathematical physics)

$$x^2 \frac{d^2 y}{dx^2} + p_1 x \frac{dy}{dx} + p_2 y = 0 \quad ((p_1, p_2) = \text{const}). \tag{18.38}$$

We put $x = e^t$, then

$$\frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = e^{-t}\frac{dy}{dt},$$

$$\frac{d^2 y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dt}\left(\frac{dy}{dx}\right)\frac{dt}{dx} = e^{-2t}\frac{d^2 y}{dt^2} - e^{-2t}\frac{dy}{dt}.$$

Substituting the expressions for x, dy/dx and d^2y/dx^2 into (18.38) we obtain the differential equation with constant coefficients

$$\frac{d^2 y}{dt^2} + (p_1 - 1)\frac{dy}{dt} + p_2 y = 0. \tag{18.39}$$

We integrate it in a conventional manner, i.e., we set up the characteristic equation

$$\lambda^2 + (p_1 - 1)\lambda + p_2 = 0, \tag{18.40}$$

find its roots and from the nature of the roots we form the general solution of (18.39). We then return to the old variable x.

Example. Find the general solution of equation

$$x^2 \frac{d^2 y}{dx^2} + 3x\frac{dy}{dx} + y = 0.$$

◄ Changing the variable $x = e^t$ yields the equation

$$\frac{d^2 y}{dt^2} + 2\frac{dy}{dt} + y = 0,$$

whose characteristic equation is

$$\lambda^2 + 2\lambda + 1 = 0.$$

It has the roots $\lambda_1 = \lambda_2 = -1$. The general solution of the transformed equation is

$$y(t) = C_1 e^{-t} + C_2 t e^{-t}.$$

In terms of x, we obtain the general solution of the original equation

$$y(x) = \frac{C_1 + C_2 \ln x}{x}, \quad x > 0. ►$$

Remark 1. For (18.39), if the roots of (18.40) are real and different, the particular solutions have the form $e^{kt} = (e^t)^k = x^k$. Therefore, from the very beginning we could seek particular solutions in this form. Substituting $y = x^k$ $(x > 0)$ into (18.38), we obtain for k the equation

$$k^2 + (p_1 - 1)k + p_2 = 0, \tag{18.41}$$

which coincides with (18.40). Corresponding to each real root of equation (18.41) is the particular solution $y = x^k$ of equation (18.38); to a double root there correspond two solutions $y_1 = x^k$ and $y_2 = x^k \ln x$ of (18.38). To the pair of complex conjugate roots $\alpha \pm i\beta$ of equation (18.41) there correspond two solutions $y_1 = x^\alpha \cos(\beta \ln x)$, $y_2 = x^\alpha \sin(\beta \ln x)$ of equation (18.38).

Remark 2. The equation

$$(ax + b)^2 \frac{d^2y}{dx^2} + p_1(ax + b)\frac{dy}{dx} + p_2 y = 0,$$

where a, b, p_1 and p_2 are constant numbers, can also be reduced to an equation with constant coefficients by the substitution $ax + b = e^t$.

18.8 Linear Inhomogeneous Differential Equations

A *linear inhomogeneous differential equation* of order n is an equation of the form

$$a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \ldots + a_n(x)y = g(x). \tag{18.42}$$

Here the functions $a_0(x)$, $a_1(x)$, ..., $a_n(x)$, $g(x)$ are defined on some interval (α, β). If $a_0(x) \neq 0$ on (α, β), then dividing by $a_0(x)$ gives

$$y^{(n)} + p_1(x)y^{(n-1)} + \ldots + p_n(x)y = f(x), \tag{18.43}$$

where

$$p_k(x) = \frac{a_k(x)}{a_0(x)}, \quad k = 1, 2, \ldots, n, \quad f(x) = \frac{g(x)}{a_0(x)}.$$

From Theorem 18.1 it follows that if on an interval $[a, b]$ the coefficients $p_k(x)$ and the right-hand side $f(x)$ of equation (18.43) are continuous, then

this equation has a unique solution subject to

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \quad \dots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)},$$

where $x_0 \in (a, b)$, $-\infty < y_0^{(k)} < +\infty$, $k = 0, 1, \dots, n-1$.

Equation (18.43) can be written as

$$L[y] = f(x), \tag{18.44}$$

where, as before, $L[y] \equiv y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y$.

Theorem 18.12. *If $\tilde{y}(x)$ is a solution of the inhomogeneous equation $L[y] = f(x)$, and $y_0(x)$ is a solution of the corresponding homogeneous equation $L[y] = 0$, then the sum $y_0(x) + \tilde{y}(x)$ is a solution of the inhomogeneous equation.*

◄ As stated, $L[\tilde{y}] \equiv f(x)$, $L[y_0] \equiv 0$. The operator L being linear, we have

$$L[y_0 + \tilde{y}] = L[y_0] + L[\tilde{y}] \equiv f(x).$$

This means that $y_0(x) + \tilde{y}(x)$ is a solution of the equation $L[y] = f(x)$. ►

Theorem 18.13. *If $y_1(x)$ is a solution of the equation $L[y] = f_1(x)$ and $y_2(x)$ is a solution of the equation $L[y] = f_2(x)$, then the function $y_1(x) + y_2(x)$ is a solution of the equation $L[y] = f_1(x) + f_2(x)$.*

◄ As stated, $L[y_1] \equiv f_1(x)$, $L[y_2] \equiv f_2(x)$. Since L is a linear operator, we get

$$L[y_1 + y_2] = L[y_1] + L[y_2] \equiv f_1(x) + f_2(x).$$

This implies that $y_1(x) + y_2(x)$ is a solution of the equation $L[y] = f_1(x) + f_2(x)$. ►

The theorem proves the so-called *superposition principle.*

Theorem 18.14. *If the equation $L[y] = U(x) + iV(x)$, where all $p_k(x)$, $k = 1, 2, \dots, n$, and $U(x)$ and $V(x)$ are real-valued, has a solution $y = u(x) + iv(x)$, then the real part of the solution $u(x)$ and its imaginary part $v(x)$ are solutions of the equations $L[y] = U(x)$ and $L[y] = V(x)$, respectively.*

◄ By the hypothesis we have

$$L[u + iv] \equiv U(x) + iV(x),$$

or

$$L[u] + iL[v] \equiv U(x) + iV(x).$$

Hence $L[u] \equiv U(x)$ and $L[v] \equiv V(x)$. ►

Theorem 18.15 (structure of general solution). *The general solution in the domain $a < x < b$, $|y^{(k)}| < +\infty$, $k = 0, 1, 2, n-1$, of the equation $L[y] \equiv y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = f(x)$ (with $p_k(x)$, $k = 1, 2, \dots, n$ and the right-hand part $f(x)$ continuous on the interval $[a, b]$, is the*

sum of the general solution of the corresponding homogeneous equation $y_{g.h.} = \sum_{i=1}^{n} C_i y_i(x)$ *and some particular solution* $y_{p.i.} = \tilde{y}(x)$ *of the inhomogeneous equation, i.e.,* $y_{g.i.} = y_{g.h.} + y_{p.i.}$.

◄ We have thus to show that

$$y(x) = \sum_{i=1}^{n} C_i y_i(x) + \tilde{y}(x), \tag{18.45}$$

where $C_1, C_2, \ldots, C_n$ are arbitrary constants, and $y_1(x), y_2(x), \ldots, y_n(x)$ are linearly independent solutions of the corresponding homogeneous equation $L[y] = 0$, is the general solution of $L[y] = f(x)$. We will rely on the definition of the general solution and simply check whether or not the family of functions $y(x)$ given by (18.45) meet the conditions (1) and (2) in that definition.

Really, the function $y(x)$ given by (18.45) is a solution of (18.43) at any constants, since the sum of a solution of an inhomogeneous equation and any solution of the corresponding homogeneous equation is a solution of the inhomogeneous equation $L[y] = f(x)$.

Since the equation (18.43) for $x \in [a, b]$ satisfies the conditions of Theorem 18.1, it only remains to show that by selecting $C_1, C_2, \ldots, C_n$ in (18.45) we can meet the arbitrary initial conditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \quad \ldots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}, \tag{18.46}$$

where $x_0 \in (a, b)$, i.e., we can solve any Cauchy problem.

We will restrict ourselves to the case of $n = 3$. If we require that (18.45) satisfy the initial conditions (18.46), we arrive at the system of equations for C_1, C_2 and C_3:

$$\begin{cases} C_1 y_1(x_0) + C_2 y_2(x_0) + C_3 y_3(x_0) + \tilde{y}(x_0) = y_0, \\ C_1 y_1'(x_0) + C_2 y_2'(x_0) + C_3 y_3'(x_0) + \tilde{y}'(x_0) = y_0', \\ C_1 y_1''(x_0) + C_2 y_2''(x_0) + C_3 y_3''(x_0) + \tilde{y}''(x_0) = y_0''. \end{cases} \tag{18.47}$$

This system of three equations linear in C_1, C_2 and C_3 permits only one solution for C_1, C_2 and C_3 for arbitrary right-hand sides, since the determinant of the system is the Wronskian $W(x_0)$ for a linearly independent system of solutions of the corresponding homogeneous equation, and hence is nonzero at any point $x \in (a, b)$, specifically at $x = x_0$. This suggests that, whatever y_0, y_0', and y_0'', we will find the solution C_1^0, C_2^0, and C_3^0 of (18.47), such that the function

$$y = C_1^0 y_1(x) + C_2^0 y_2(x) + C_3^0 y_3(x) + \tilde{y}(x)$$

will be a solution of the differential equation (18.43), obeying the initial conditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \quad y''|_{x=x_0} = y_0''. ►$$

It follows from the theorem that the problem of finding the general solution of a linear inhomogeneous equation consists in finding some particular solution of this inhomogeneous equation and the general solution of the corresponding homogeneous equation.

Example. Find the general solution of the equation $y'' + y = x$.
◄ It is easily seen that the function $\tilde{y}(x) = x$ is a particular solution of the equation. To find the general solution we only have to find the general solution of the corresponding homogeneous equation

$$y'' + y = 0. \tag{*}$$

This is a linear homogeneous equation with constant coefficients. The appropriate characteristic equation is $\lambda^2 + 1 = 0$, its roots are $\lambda_1 = i$ and $\lambda_2 = -i$. Therefore, the general solution of (*) has the form

$$y_{\text{g.h.}} = C_1 \cos x + C_2 \sin x.$$

The general solution of the original inhomogeneous equation is

$$y_{\text{g.i.}} = C_1 \cos x + C_2 \sin x + x. ►$$

18.9 Integration of Linear Inhomogeneous Equation by Variation of Constants

For simplicity we will start with the case of the second-order equation:

$$y'' + p_1(x)y' + p_2(x)y = f(x), \tag{18.48}$$

where $p_1(x)$, $p_2(x)$, and $f(x)$ are continuous on the interval $[a, b]$. Suppose that we know the fundamental set of solutions $y_1(x)$ and $y_2(x)$ of the corresponding homogeneous equation

$$y'' + p_1(x)y' + p_2(x)y = 0, \tag{18.49}$$

then

$$y = C_1 y_1(x) + C_2 y_2(x),$$

where C_1 and C_2 are constants, is the general solution of (18.49). Notice that this assumption is a bit uncomfortable, since there is no general method of finding solutions of linear homogeneous equations of order $n \geqslant 2$ with variable coefficients.

We will integrate (18.48) using the method of variation of constants (Lagrange's method). We will seek the solution of (18.48) in the form

$$y = C_1(x)y_1(x) + C_2(x)y_2(x), \tag{18.50}$$

where $C_1(x)$ and $C_2(x)$ are new unknown functions of x. To find them we will need two equations containing them. Clearly, $C_1(x)$ and $C_2(x)$ must obey the equation that results from the substitution of $C_1(x)y_1(x) + C_2(x)y_2(x)$ for $y(x)$ into the original equation.

We now subject $C_1(x)$ and $C_2(x)$ to additional conditions. Differentiating (18.50) gives

$$y' = C_1(x)y_1' + C_2(x)y_2' + C_1'(x)y_1 + C_2'(x)y_2.$$

The additional condition (its expediency will be obvious from the following) is

$$C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0, \tag{18.51}$$

then

$$y' = C_1(x)y_1' + C_2(x)y_2', \tag{18.52}$$

$$y'' = C_1y_1'' + C_2y_2'' + C_1'y_1' + C_2'y_2'. \tag{18.53}$$

Substituting (18.50), (18.52) and (18.53) for y, y' and y'' in (18.48) gives

$$C_1(x)[y_1'' + p_1(x)y_1' + p_2(x)y_1] + C_2(x)[y_2'' + p_1(x)y_2' + p_2(x)y_2] + C_1'(x)y_1' + C_2'(x)y_2' = f(x).$$

The expressions in brackets are identically zero, since $y_1(x)$ and $y_2(x)$ are solutions of the homogeneous equation (18.49). And so substituting y, y' and y'' into (18.48) we obtain

$$C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = f(x). \tag{18.54}$$

Thus, $y = C_1(x)y_1(x) + C_2(x)y_2(x)$ will be a solution of the inhomogeneous differential equation (18.48), if $C_1(x)$ and $C_2(x)$ simultaneously obey (18.51) and (18.54), i.e., the system

$$\begin{cases} C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0, \\ C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = f(x), \end{cases} \tag{18.55}$$

whose determinant is the Wronskian of the linearly independent solutions $y_1(x)$ and $y_2(x)$ of (18.49) and, therefore, nonzero everywhere in the interval (a, b). We solve the system as a linear algebraic system for $C_1'(x)$ and $C_2'(x)$

$$C_1'(x) = \varphi_1(x), \quad C_2'(x) = \varphi_2(x)$$

(here $\varphi_1(x)$ and $\varphi_2(x)$ are known functions) and integrate

$$C_1(x) = \int \varphi_1(x)\,dx + C_1, \quad C_2(x) = \int \varphi_2(x)\,dx + C_2$$

(here C_1 and C_2 are constants of integration). Substituting these expressions for $C_1(x)$ amd $C_2(x)$ into (18.50), we find the general solution of (18.48)

$$y = C_1 y_1(x) + C_2 y_2(x) + y_1(x)\int \varphi_1(x)\,dx + y_2(x)\int \varphi_2(x)\,dx,$$

where C_1 and C_2 are arbitrary constants.

To sum up: given a fundamental set of equations of the corresponding homogeneous equation, we can find the general solution of an inhomogeneous equation by quadratures.

Example. Find the general solution of the equation

$$y'' + y = \frac{1}{\sin x}.$$

◄ Consider the homogeneous equation $y'' + y = 0$, which corresponds to the original inhomogeneous equation. This is an equation with constant coefficients. The functions $y_1(x) = \sin x$ and $y_2(x) = \cos x$ form its fundamental set of solutions. We will look for the solution of the original equation in the form

$$y = C_1(x)\sin x + C_2(x)\cos x. \tag{$*$}$$

System (18.55) will then become

$$\begin{cases} C_1'(x)\sin x + C_2'(x)\cos x = 0, \\ C_1'(x)\cos x - C_2'(x)\sin x = \dfrac{1}{\sin x}. \end{cases}$$

Solving the system for $C_1'(x)$ and $C_2'(x)$ gives

$$C_1'(x) = \frac{\cos x}{\sin x}, \qquad C_2'(x) = -1,$$

$$C_1(x) = \ln|\sin x| + C_1, \qquad C_2(x) = -x + C_2.$$

Substituting these expressions into (*), we will find the general solution of the original equation

$$y = C_1\sin x + C_2\cos x + \sin x \ln|\sin x| - x\cos x. \text{ ►}$$

To integrate a linear inhomogeneous differential equation of order n $(n \geqslant 1)$

$$y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = f(x) \tag{18.56}$$

we will proceed in a similar manner.

Let $y_1(x), y_2(x), \dots, y_n(x)$ be a known fundamental set of solutions of the corresponding homogeneous equation. We will seek the solution of

(18.56) in the form

$$y = C_1(x)\,y_1(x) + C_2(x)\,y_2(x) + \ldots + C_n(x)\,y_n(x), \tag{18.57}$$

where $C_1(x)$, $C_2(x)$, ..., $C_n(x)$ are new unknown functions.

To find the n functions $C_1(x)$, $C_2(x)$, ..., $C_n(x)$, we will form a system of n equations that contain these functions. When we set up the system, we can take $n - 1$ equations in an arbitrary manner and then form the nth equation proceeding from the requirement that $y(x)$ given by (18.57) obeyed (18.56). The first $n - 1$ equations will be

$$C_1'(x)\,y_1 + C_2'(x)\,y_2 + \ldots + C_n'(x)\,y_n = 0,$$
$$C_1'(x)\,y_1' + C_2'(x)\,y_2' + \ldots + C_n'(x)\,y_n' = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$C_1'(x)\,y_1^{(n-2)} + C_2'(x)\,y_2^{(n-2)} + \ldots + C_n'(x)\,y_n^{(n-2)} = 0.$$

Then, for the function $y(x)$ given by (18.57) to satisfy (18.56) $C_i(x)$ must be subject to the condition

$$C_1'(x)\,y_1^{(n-1)} + C_2'(x)\,y_2^{(n-1)} + \ldots + C_n'(x)\,y_n^{(n-1)} = f(x).$$

To find $C_i'(x)$, $i = 1, 2, \ldots, n$, we will have the system

$$\begin{cases} C_1'y_1 + C_2'y_2 + \ldots + C_n'y_n = 0, \\ C_1'y_1' + C_2'y_2' + \ldots + C_n'y_n' = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ C_1'y_1^{(n-2)} + C_2'y_2^{(n-2)} + \ldots + C_n'y_n^{(n-2)} = 0, \\ C_1'y_1^{(n-1)} + C_2'y_2^{(n-1)} + \ldots + C_n'y_n^{(n-1)} = f(x). \end{cases} \tag{18.58}$$

The determinant of the system is the Wronskian of a fundamental set of solutions of a homogeneous equation, and hence it is nonzero in (a, b). Therefore, system (18.58) is solved uniquely in $C_i'(x)$, $i = 1, 2, \ldots, n$. Solving it gives $C_i'(x) = \varphi_i(x)$, where $\varphi_i(x)$ are known functions, and then

$$C_i(x) = \int \varphi_i(x)\,dx + C_i \quad (i = 1, 2, \ldots, n).$$

Substituting these into (18.57), we will obtain the general solution $y(x)$ of the original equation (18.56)

$$y(x) = \sum_{i=1}^{n} C_i y_i(x) + \sum_{i=1}^{n} y_i(x) \int \varphi_i(x)\,dx,$$

where C_1, C_2, ..., C_n are arbitrary constants.

18.10 Inhomogeneous Linear Differential Equations with Constant Coefficients

In Section 18.9 we discussed the general method of solving an inhomogeneous linear differential equation, namely the method of variation of constants. When the equation has constant coefficients, a particular solution of the inhomogeneous equation can sometimes be found much easier, namely by selection.

Consider some forms of equations to which the method can be applied.

(1) Equations of the type

$$L[y] \equiv y^{(n)} + p_1 y^{(n-1)} + \dots + p_n y = P_m(x), \tag{18.59}$$

where $p_1, p_2, p_3, \dots, p_n$ are real numbers, $P_m(x)$ is a given polynomial of degree m:

$$P_m(x) = A_0 x^m + A_1 x^{m-1} + \dots + A_m, \quad A_0 \neq 0.$$

The characteristic equation for the homogeneous equation that corresponds to (18.59) has the form

$$\lambda^n + p_1 \lambda^{n-1} + \dots + p_{n-1}\lambda + p_n = 0 \quad \text{or} \quad \varphi(\lambda) = 0,$$

where $\varphi(\lambda) \equiv \lambda^n + p_1\lambda^{n-1} + \dots + p_n$ is the characteristic polynomial.

If p_n is nonzero, i.e., $\lambda = 0$ is no root of $\varphi(\lambda) = 0$, then there exists a particular solution $\tilde{y}(x)$ of (18.59) that also has the form of a polynomial of degree m. Actually, we take $\tilde{y}(x)$ in the form

$$\tilde{y}(x) = B_0 x^m + B_1 x^{m-1} + \dots + B_m,$$

where $B_0, B_1, \dots, B_m$ are indefinite coefficients. We then substitute this into (18.59) and, comparing the coefficients at the same powers of x on the left- and right-hand sides, we derive for B_i a system of linear algebraic equations, which is always solvable, if $p_n \neq 0$. Indeed, equating the coefficients at x^m, x^{m-1}, ..., we obtain

$$\begin{array}{l|lll}
x^m & p_n B_0 = A_0, & \text{hence} & B_0 = \dfrac{A_0}{p_n}, \\[2ex]
x^{m-1} & m p_{n-1} B_0 + p_n B_1 = A_1, & \text{hence} & B_1 = \dfrac{A_1 - m p_{n-1} B_0}{p_n}.
\end{array}$$

In summary, if $\lambda = 0$ is no root of the characteristic equation $\varphi(\lambda) = 0$, then there exists a particular solution $\tilde{y}(x)$ of (18.59) which has the form of a polynomial whose degree equals that of a polynomial on the right of (18.59):

$$\tilde{y}(x) = B_0 x^m + B_1 x^{m-1} + \dots + B_m. \tag{18.60}$$

Suppose now that $p_n = 0$. To be more general, we also assume that $p_{n-1} = p_{n-2} = \ldots = p_{n-r+1} = 0$, but $p_{n-r} \neq 0$, i.e., $\lambda = 0$ is an r-tuple root ($r \geqslant 1$) of the characteristic equation $\varphi(\lambda) = 0$. Equation (18.59) then becomes

$$y^{(n)} + p_1 y^{(n-1)} + \ldots + p_{n-r} y^{(r)} = P_m(x). \tag{18.61}$$

Putting $y^{(r)} = z$, we arrive at the previous case. Therefore, there exists a particular solution of (18.61) in the form

$$\tilde{y}^{(r)}(x) = \tilde{B}_0 x^m + \tilde{B}_1 x^{m-1} + \ldots + \tilde{B}_m.$$

We deduce from this that $\tilde{y}(x)$ is a polynomial of degree $m + r$, such that its terms with x to power $r - 1$ and lower will have arbitrary constant coefficients, which can, in particular, be chosen to be zero. The particular solution will then be

$$\tilde{y}(x) = x^r(B_0 x^m + B_1 x^{m-1} + \ldots + B_m).$$

To sum up: if $\lambda = 0$ is a root of multiplicity $r \geqslant 1$ of the characteristic equation $\varphi(\lambda) = 0$, then the particular solution $\tilde{y}(x)$ of (18.59) should be sought for in the form of the product of x^r by the polynomial $Q_m(x)$ of degree m with indefinite coefficients: $\tilde{y}(x) = x^r Q_m(x)$.

Example. Find a particular solution of the equation

$$y'' + y' = 2x + 3.$$

◄ The characteristic equation $\lambda^2 + \lambda = 0$ has roots $\lambda_1 = 0$ and $\lambda_2 = -1$, therefore $\lambda = 0$ is a simple root ($r = 1$) of the equation. The polynomial on the right-hand side is a polynomial of the first degree ($m = 1$), therefore we will seek the particular solution of the inhomogeneous differential equation as

$$\tilde{y}(x) = x(B_0 x + B_1).$$

Substituting $\tilde{y}(x)$ into the equation and comparing the coefficients at the same degrees of x, we will find that $B_0 = 1$ and $B_1 = 1$. And so the desired particular solution will be $\tilde{y}(x) = x^2 + x$. ►

(2) Equations of the type

$$L[y] \equiv y^{(n)} + p_1 y^{(n-1)} + \ldots + p_n y = e^{ax} P_m(x), \quad a = \text{const}. \tag{18.62}$$

We will seek the particular solution in the form

$$\tilde{y}(x) = e^{ax} z(x),$$

where $z = z(x)$ is to be found from the condition

$$L[e^{ax} z] \equiv e^{ax} P_m(x). \tag{18.63}$$

Then

$$\tilde{y} = e^{ax}z,$$
$$\tilde{y}' = e^{ax}(az + z'),$$
$$\dots\dots\dots\dots\dots\dots$$
$$\tilde{y}^{(r)} = e^{ax}\left(a^r z + \frac{r}{1!}a^{r-1}z' + \frac{r(r-1)}{2!}a^{r-2}z'' + \dots + z^{(r)}\right),$$
$$\dots\dots\dots\dots\dots\dots$$
$$\tilde{y}^{(n)} = e^{ax}\left(a^n z + na^{n-1}z' + \frac{n(n-1)}{2!}a^{n-2}z'' + \dots + z^{(n)}\right).$$

We then multiply $\tilde{y}, \tilde{y}', \dots, \tilde{y}^{(n)}$ by $p_n, p_{n-1}, \dots, 1$, respectively, and add up the results, grouping the terms by columns:

$$L[e^{ax}z] = e^{ax}\left\{\varphi(a)z + \frac{\varphi'(a)}{1!}z' + \dots + \frac{\varphi^{(r)}(a)}{r!}z^{(r)} + \dots + z^{(n)}\right\}.$$

Here $\varphi(a)$ results from substituting $\lambda = a$ in the characteristic polynomial $\varphi(\lambda)$. It follows that to obtain the identity (18.63) we should define $z(x)$ as a solution of

$$z^{(n)} + \dots + \frac{\varphi^{(r)}(a)}{r!}z^{(r)} + \dots + \varphi'(a)z' + \varphi(a)z = P_m(x). \quad (18.64)$$

This is a linearly inhomogeneous equation with constant coefficients, its right-hand side being a polynomial. Therefore, we should seek a particular solution of (18.64) as a polynomial $Q_m(x)$ of degree m, if $\varphi(a) \neq 0$, i.e., when a is not a root of the characteristic equation $\varphi(\lambda) = 0$. If a is a root of the characteristic equation of multiplicity $r \geqslant 1$, then

$$\varphi(a) = \varphi'(a) = \dots = \varphi^{(r-1)}(a) = 0, \quad \varphi^{(r)}(a) \neq 0,$$

and we should seek the solution of (18.64) as $x^r Q_m(x)$. Therefore we will look for the particular solution of the original equation (18.62) in the form

$$\tilde{y}(x) = e^{ax}Q_m(x), \quad (18.65)$$

if a is not a root of the characteristic equation $\varphi(\lambda) = 0$, and in the form

$$\tilde{y}(x) = x^r e^{ax}Q_m(x), \quad (18.66)$$

if a is a root of the characteristic equation of multiplicity $r \geqslant 1$. Here $Q_m(x)$ is a polynomial of degree m with indefinite coefficients, i.e.,

$$Q_m(x) = B_0x^m + B_1x^{m-1} + \dots + B_m.$$

Examples. (1) Find the particular solution of the equation $y'' + y' = 2e^x$.

◄ The characteristic equation $\lambda^2 + \lambda = 0$ has the roots $\lambda_1 = 0$ and $\lambda_2 = -1$. The right-hand side of the equation is the product of e^x $(a = 1)$ by a zero-degree polynomial $(m = 0)$. Since $a = 1$ is not a root of the characteristic equation, the particular solution $\tilde{y}(x)$ should be sought for in the form

$$\tilde{y}(x) = Be^x.$$

Substituting $\tilde{y}(x)$ into the equation and reducing by e^x, we will find $B = 1$. Hence $\tilde{y}(x) = e^x$.

(2) Determine the form of the particular solution of the equation $y'' - 2y' + y = xe^x$.

◄ The characteristic equation $\lambda^2 - 2\lambda + 1 = 0$ has the roots $\lambda_1 = \lambda_2 = 1$. In this case $P_m(x) \equiv x$, i.e., $m = 1$ and $a = 1$ is a double root $(r = 2)$ of the characteristic equation. Therefore, the particular solution $\tilde{y}(x)$ should be sought for in the form

$$\tilde{y}(x) = x^2e^x(B_0x + B_1). \text{ ►}$$

(3) The above reasoning also holds true for complex a. Therefore, if the right-hand side of the linear differential equation

$$L[y] \equiv y^{(n)} + p_1y^{(n-1)} + \dots + p_ny = f(x) \tag{18.67}$$

has the form

$$f(x) = e^{\alpha x}[P_m(x)\cos\beta x + Q_s(x)\sin\beta x], \tag{18.68}$$

where $P_m(x)$ and $Q_s(x)$ are polynomials of degrees m and s, respectively, then we proceed as follows. We transform the trigonometric functions into exponential functions using the Euler formulas

$$\cos\beta x = \frac{e^{i\beta x} + e^{-i\beta x}}{2}, \quad \sin\beta x = \frac{e^{i\beta x} - e^{-i\beta x}}{2i},$$

then

$$\begin{aligned} f(x) &= e^{\alpha x}P_m(x)\frac{e^{i\beta x} + e^{-i\beta x}}{2} + e^{\alpha x}Q_s(x)\frac{e^{i\beta x} - e^{-i\beta x}}{2i} \\ &= \left[\frac{1}{2}P_m(x) + \frac{1}{2i}Q_s(x)\right]e^{(\alpha + i\beta)x} \\ &+ \left[\frac{1}{2}P_m(x) - \frac{1}{2i}Q_s(x)\right]e^{(\alpha - i\beta)x}. \end{aligned}$$

The polynomials in brackets have a degree equal to the highest degree of $P_m(x)$ and $Q_s(x)$. We denote the polynomials by $M(x)$ and $N(x)$ to obtain

on the right

$$M(x)\,e^{(\alpha + i\beta)x} + N(x)\,e^{(\alpha - i\beta)x}. \tag{18.69}$$

For each term on the right we can use the rule mentioned above: if $\alpha \pm i\beta$ are no roots of a characteristic equation, then a particular solution of the differential equation can be sought for in the form (18.69). If then $\alpha \pm i\beta$ are roots of a characteristic equation of multiplicity $r \geqslant 1$, then the particular solution also includes the factor x^r.

If we return to the trigonometric functions, the rule could be formulated as follows:

(a) if $\alpha \pm i\beta$ are no roots of a characteristic equation, then a particular solution $\tilde{y}(x)$ of the differential equation (18.67) should be sought in the form

$$\tilde{y}(x) = e^{\alpha x}[U(x)\cos\beta x + V(x)\sin\beta x], \tag{18.70}$$

where $U(x)$ and $V(x)$ are polynomials with indefinite coefficients each of which has a degree equal to the highest of the degrees of $P_m(x)$ and $Q_s(x)$. The coefficients of these polynomials are found by substituting $\tilde{y}(x)$ into the differential equation and equating the coefficients at the same powers of x on either side. We should equate to each other the corresponding coefficients of the polynomials that appear as factors at $\cos\beta x$, and separately the coefficients at $\sin\beta x$;

(b) if $\alpha \pm i\beta$ are r-tuple roots of the characteristic equation (resonance), then the particular solution $\tilde{y}(x)$ must be sought as

$$\tilde{y}(x) = x^r e^{\alpha x}[U(x)\cos\beta x + V(x)\sin\beta x]. \tag{18.71}$$

Remark. The forms of particular solutions (18.70) and (18.71) are also valid when on the right-hand side of the equation either $Q_s(x)$ or $P_m(x)$ is identically zero, i.e., when the right-hand side has the form

$$e^{\alpha x}P_m(x)\cos\beta x \quad \text{or} \quad e^{\alpha x}Q_s(x)\sin\beta x.$$

Examples. (1) Find a particular solution of the equation $y'' + 2y' + y = \cos x$.

◄ The characteristic equation $\lambda^2 + 2\lambda + 1 = 0$ has the roots $\lambda_1 = \lambda_2 = -1$. In this case $\alpha = 0$ and $\beta = 1$, therefore $\alpha \pm i\beta = \pm i$ are no roots of the characteristic equation. Further, $P_m(x) \equiv 1$ and $Q_s(x) \equiv 0$, and so the particular solution of the equation must be sought for in the form

$$\tilde{y}(x) = A\cos x + B\sin x, \quad A,\ B = \text{const.}$$

Substituting $\tilde{y}(x)$ into the equation, we will have $A = 0$ and $B = 1/2$, and so $\tilde{y}(x) = (\sin x)/2$. ►

(2) (*Resonance.*) Consider the equation of elastic vibrations without resistance under the action of a periodic external force

$$y'' + \omega^2 y = a \sin \beta t, \qquad (*)$$

where the independent variable is the time t.
◄ The general solution of the homogeneous equation is

$$y_{\text{g.h.}} = C_1 \cos \omega t + C_2 \sin \omega t = A \sin (\omega t + \delta).$$

If $\beta \neq \omega$, i.e., if the frequency of the external force does not coincide with the frequency ω of the natural oscillations of the system, then the particular solution of the inhomogeneous equation will be

$$\tilde{y}(t) = M \cos \beta t + N \sin \beta t, \quad M, N = \text{const}.$$

Substituting this expression into (*), we will find that $M = 0$ and $N = a/(\omega^2 - \beta^2)$. The general solution of (*) will then be

$$y_{\text{g.i.}} = A \sin (\omega t + \delta) + \frac{a}{\omega^2 - \beta^2} \sin \beta t,$$

i.e., the resultant motion is a combination of natural oscillations with frequency ω and forced oscillations with frequency β.

If $\beta = \omega$, i.e., the frequency of the external force coincides with the frequency of natural oscillations of the system, then we have to seek the particular solution of the inhomogeneous equation (*) in the form

$$\tilde{y}(t) = t(M \cos \omega t + N \sin \omega t).$$

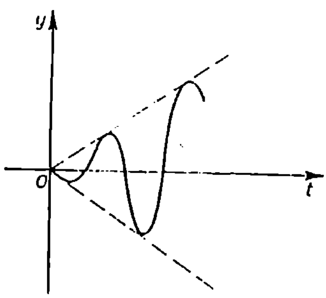


Fig. 18.3

Substituting $\tilde{y}(t)$ into (*) gives

$$M = -\frac{a}{2\omega}, \quad N = 0, \quad \text{hence} \quad \tilde{y}(t) = -\frac{a}{2\omega} t \cos \omega t.$$

The general solution of (*) will be

$$y = A \sin(\omega t + \delta) - \frac{a}{2\omega} t \cos \omega t. \tag{**}$$

The second term on the right of (**) suggests that here the amplitude of oscillations increases indefinitely with t (Fig. 18.3). This phenomenon, which occurs when the frequency of the external force coincides with the frequency of natural oscillations of the system, is called the *resonance.* ►

It is convenient to look for particular solutions in the following manner. Consider the linear inhomogeneous differential equation with real constant coefficients

$$L[y] \equiv y^{(n)} + p_1 y^{(n-1)} + \dots + p_n y = e^{\alpha x} P_m(x) \cos \beta x, \tag{18.72}$$

where $P_m(x)$ is a specified polynomial of degree m with real coefficients; α and β are real numbers.

We now set up an auxiliary inhomogeneous equation with the same left-hand side as (18.72) and the right-hand side represented as a complex-valued function of a real variable x

$$\begin{aligned} L[z] &\equiv z^{(n)} + p_1 z^{(n-1)} + \dots + p_n z \\ &= e^{\alpha x} P_m(x) \cos \beta x + i e^{\alpha x} P_m(x) \sin \beta x. \end{aligned} \tag{18.73}$$

The right-hand side of (18.72) is the real part of the right-hand side of (18.73); therefore, by virtue of Theorem 18.14, the real part $u(x)$ of the solution $z(x) = u(x) + iv(x)$ of equation (18.73) will be a solution of the original equation (18.72). The problem thus comes down to finding a particular solution of (18.73), which can be written as

$$L[z] = e^{(\alpha + i\beta)x} P_m(x). \tag{18.74}$$

It follows from the foregoing that:

(1) if the number $\alpha + i\beta$ is not a root of the characteristic equation $\varphi(\lambda) \equiv \lambda^n + p_1\lambda^{n-1} + \dots + p_n = 0$, then the particular solution of equation (18.74) should be sought for in the form

$$\tilde{z}(x) = e^{(\alpha + i\beta)x} Q_m(x), \tag{18.75}$$

where

$$Q_m(x) = B_0 x^m + B_1 x^{m-1} + \dots + B_m$$

is a polynomial of degree m with indefinite coefficients;

(2) if $\alpha + i\beta$ is a root of multiplicity r of the characteristic equation, then the particular solution of (18.74) has the form

$$\tilde{z}(x) = x^r e^{(\alpha + i\beta)x} Q_m(x). \tag{18.76}$$

The replacement of trigonometric functions by exponential functions simplifies the calculations, since the substitution of $\tilde{z}(x)$ into (18.74) makes it possible to cancel out $e^{(\alpha + i\beta)x}$ on either side. The complex coefficients $B_0, B_1, \ldots, B_m$ in $Q_m(x)$ are found by substituting solutions (18.75) or (18.76) into (18.74) and equating the coefficients at the same powers of x on either side. We then isolate the real part $u(x)$ of solutions (18.75) and (18.76) to find the particular solution of (18.72).

If the equation has the form

$$L[y] = e^{\alpha x} P_m(x) \sin \beta x, \tag{18.77}$$

we proceed in a similar manner: (1) we pass on to the auxiliary equation (18.74); (2) we then find the particular solution $\tilde{z}(x) = u(x) + iv(x)$ of the equation. The imaginary part $v(x)$ of the solution will be a particular solution of equation (18.77).

Example. Find the particular solution of the equation

$$y'' + y = 5xe^{-x} \cos x. \tag{*}$$

◄ We write the auxiliary equation

$$z'' + z = 5xe^{(-1+i)x}. \tag{**}$$

Since the number $-1 + i$ is not a root of the characteristic equation $\lambda^2 + 1 = 0$, the particular solution of (**) will be sought for as

$$\tilde{z}(x) = (B_0 x + B_1) e^{(-1+i)x}.$$

Substituting $z(x)$ and $\tilde{z}''(x) = 2(-1 + i) B_0 e^{(-1+i)x} + (-1 + i)^2 (B_0 x + B_1) e^{(-1+i)x}$ into equation (**) and cancelling out $e^{(-1+i)x}$, we obtain

$$(-2 + 2i) B_0 + (-2i + 1)(B_0 x + B_1) = 5x.$$

Equating the coefficients at the same powers of x on either side of this equality, we get

$$B_0(1 - 2i) = 5, \quad \text{hence} \quad B_0 = 1 + 2i;$$

$$B_1(1 - 2i) + (-2 + 2i) B_0 = 0, \quad \text{hence} \quad B_1 = \frac{2 + 14i}{5}.$$

Therefore,

$$\tilde{z}(x) = \left[(1 + 2i)x + \frac{2}{5} + \frac{14}{5}i\right] e^{-x}(\cos x + i \sin x).$$

The particular solution of the original equation will thus be

$$u(x) = \operatorname{Re} \tilde{z}(x) = e^{-x}\left[\left(x + \frac{2}{5}\right)\cos x - \left(2x + \frac{14}{5}\right)\sin x\right]. \ ►$$

18.11 Integration of Differential Equations Using Power Series and Generalized Power Series

Consider the differential equation

$$y'' = f(x, y, y'). \tag{18.78}$$

We would like to find a solution of the equation subject to the initial conditions

$$y(x_0) = y_0, \quad y'(x_0) = y_0'. \tag{18.79}$$

Suppose that the function f is analytic in the neighbourhood of the point (x_0, y_0, y_0'), i.e., is representable by a power series in $x - x_0$, $y - y_0$, and $y' - y_0'$.

The solution $y(x)$ of the Cauchy problem (18.78)-(18.79) can then be obtained as a series:

$$y(x) = y(x_0) + \frac{y'(x_0)}{1!}(x - x_0) + \frac{y''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{y^{(n)}(x_0)}{n!}(x - x_0)^n + \ldots . \tag{18.80}$$

In fact, knowing x_0, y_0, y_0', we by (18.78) itself will find $y''(x_0) = f(x_0, y_0, y_0')$. Differentiating (18.78) with respect to x gives

$$y''' = f_x' + f_y' y' + f_{y'}' y'' \tag{18.81}$$

Substituting into the right-hand side of (18.81) the values x_0, y_0 and y_0' and the value $y''(x_0)$ we have just found, we will have $y'''(x_0)$, and so on.

If series (18.80) converges in an interval $(x_0 - h, x_0 + h)$, then it defines there the solution of the Cauchy problem (18.78)-(18.79).

Example. Find the solution of the Cauchy problem

$$y'' = y, \tag{*}$$

$$y(0) = 1, \quad y'(0) = 1. \tag{**}$$

◄ By (*) and (**) we have $y''(0) = 1$. Differentiating (*) gives $y''' = y'$, and so $y'''(0) = y'(0) = 1$. In general $y^{(n)}(0) = 1$, $n = 0, 1, 2, \ldots$. Then

$$y(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots = e^x. \text{ ►}$$

Consider the linear differential equaton

$$p_0(x)y'' + p_1(x)y' + p_2(x)y = 0. \tag{18.82}$$

Theorem 18.16 (analyticity of solution). *If $p_0(x)$, $p_1(x)$ and $p_2(x)$ are analytic functions in the neighbourhood of the point $x = x_0$ and $p(x_0) \neq 0$, then the solutions of* (18.82) *are also analytic functions in a certain neighbourhood of $x = x_0$, and so these solutions can be sought for in the form of the sum of the series*

$$y = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots + a_n(x - x_0)^n + \ldots . \quad (18.83)$$

Example. Find the solution of the problem

$$y'' + y = 0, \quad y(0) = 0, \quad y'(0) = 1.$$

◄ We will look for the solution in the form of the series

$$y(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots + a_nx^n + \ldots .$$

We thus have

$$y'(x) = a_1 + 2a_2x + 3a_3x^2 + \ldots + na_nx^{n-1} + \ldots,$$
$$y''(x) = 1 \cdot 2a_2 + 2 \cdot 3a_3x + \ldots + n(n-1)\,a_nx^{n-2} + \ldots .$$

Substituting $y(x)$ and $y''(x)$ into the original equation and equating to zero the coefficients at the powers of x, we find

x^0	$a_0 + 1 \times 2a_2 = 0,$	hence $a_2 = -\dfrac{a_0}{1 \times 2},$
x	$a_1 + 2 \times 3a_3 = 0,$	hence $a_3 = -\dfrac{a_1}{2 \times 3},$
$\vdots$	....	....
x^{n-2}	$a_{n-2} + n(n-1)\,a_n = 0,$	hence $a_n = -\dfrac{a_{n-2}}{(n-1)\,n},$
	....	....

From the initial conditions we have $a_0 = 0$, $a_1 = 1$, therefore $a_2 = 0$ and, in general, $a_{2k} = 0$, $k = 1, 2, \ldots$. Further, we have $a_1 = 1$, $a_3 = \dfrac{-1}{2 \times 3}$, $a_5 = \dfrac{-a_3}{4 \times 5} = \dfrac{1}{1 \times 2 \times 3 \times 4 \times 5}$. In general, $a_{2k+1} = \dfrac{(-1)^k}{(2k+1)!}$.

We arrive at

$$y(x) = \frac{x}{1!} - \frac{x^3}{3} + \ldots + (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \ldots = \sin x. \ ►$$

Suppose now that the coefficient $p_0(x)$ becomes zero at a point x_0.

Definition. A point x_0 is called a *zero of order (multiplicity) m* (m is a positive number) of the function $f(x)$, if $f(x)$ is represented in the form $f(x) = (x - x_0)^m f_1(x)$, where $f_1(x_0) \neq 0$.

Theorem 18.17 (expandability of solution into generalized power series). *If in* (18.82) *the coefficients* $p_0(x)$, $p_1(x)$, *and* $p_2(x)$ *are analytic functions in the neighbourhood of* x_0, *where* $x = x_0$ *is a zero of order* m *of* $p_0(x)$, *a zero of order* $m - 1$ *or higher of* $p_1(x)$ (if $m > 1$), *and a zero of order* $m - 2$ *or higher of* $p_2(x)$ (if $m > 2$), *then there exists at least one nontrivial solution of* (18.82) *in the form of the sum of the generalized power series*

$$y(x) = a_0(x - x_0)^\sigma + a_1(x - x_0)^{\sigma+1} + \ldots + a_n(x - x_0)^{n+\sigma} + \ldots,$$

where σ *is a real number, not necessarily an integer.*

18.12 Bessel Equation. Bessel Functions

An equation of the form

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0, \tag{18.84}$$

where ν is a real number, is called the *Bessel equation.* This equation has a singularity at $x = 0$ (the coefficient at the highest derivative in the equation vanishes at $x = 0$). Comparing (18.82) and (18.84) indicates that for the Bessel equation $p_0(x) = x^2$, $p_1(x) = x$, $p_2(x) = x^2 - \nu^2$, since $x = 0$ is a zero of the second order ($m = 2$) of the function $p_0(x)$, is a zero of the first order of the function $p_1(x)$, and is no zero of the function $p_2(x)$ (if $\nu \neq 0$). Therefore, by virtue of Theorem 18.17, there exists a solution of (18.84) in the form of the sum of the generalized power series

$$y(x) = x^\sigma \sum_{k=0}^{\infty} a_k x^k, \quad a_0 \neq 0, \tag{18.85}$$

where σ is the *characteristic exponent* to be determined.

We rewrite (18.85) in the form

$$y(x) = \sum_{k=0}^{\infty} a_k x^{k+\sigma}$$

and find the derivatives

$$y' = \sum_{k=0}^{\infty} (k + \sigma)\, a_k x^{k+\sigma-1},$$

$$y'' = \sum_{k=0}^{\infty} (k + \sigma)(k + \sigma - 1)\, a_k x^{k+\sigma-2}.$$

Substituting these into (18.84) gives

$$x^2 \sum_{k=0}^{\infty} (k + \sigma)(k + \sigma - 1)\, a_k x^{k+\sigma-2}$$
$$+ x \sum_{k=0}^{\infty} (k + \sigma)\, a_k x^{k+\sigma-1} + (x^2 - \nu^2) \sum_{k=0}^{\infty} a_k x^{k+\sigma} = 0.$$

If we then equate to zero the coefficients at x^{σ}, $x^{\sigma+1}$, ..., $x^{\sigma+k}$, ..., we will get the system of equations

$$
\begin{array}{l|l}
x^{\sigma} & [\sigma^2 - \nu^2]\, a_0 = 0, \\
x^{\sigma+1} & [(\sigma + 1)^2 - \nu^2]\, a_1 = 0, \\
x^{\sigma+2} & [(\sigma + 2)^2 - \nu^2]\, a_2 + a_0 = 0, \qquad (k = 2, 3, \dots) \qquad (18.86) \\
\vdots & \dots\dots\dots\dots\dots\dots\dots\dots \\
x^{\sigma+k} & [(\sigma + k)^2 - \nu^2]\, a_k + a_{k-2} = 0, \\
\vdots & \dots\dots\dots\dots\dots\dots\dots\dots
\end{array}
$$

Since $a_0 \neq 0$, then from the first of (18.86) it follows that $\sigma^2 - \nu^2 = 0$, or

$$\sigma = \pm \nu.$$

Now from the second of (18.86) we will have

$$a_1 = 0.$$

We will first take the case of $\sigma = \nu > 0$. We rewrite the kth ($k > 1$) equation in (18.86) in the form

$$(\sigma + k + \nu)(\sigma + k - \nu)\, a_k + a_{k-2} = 0.$$

From this we derive the recurrence formula to determine a_k in terms of a_{k-2}

$$a_k = -\frac{a_{k-2}}{(\sigma + k + \nu)(\sigma + k - \nu)}.$$

Considering that $a_1 = 0$, we obtain from this that $a_3 = 0$ and, in general, $a_{2m+1} = 0$. On the other hand, each even coefficient can be expressed through the previous one by the formula

$$a_{2m} = -\frac{a_{2m-2}}{(\sigma + 2m + \nu)(\sigma + 2m - \nu)},$$

or, since $\sigma = \nu$,

$$a_{2m} = -\frac{a_{2m-2}}{2^2 m(m + \nu)}.$$

If we apply this formula several times we will be able to express a_{2m} through a_0

$$a_2 = -\frac{a_0}{2^2 \times 1 \times (\nu + 1)},$$

$$a_4 = -\frac{a_2}{2^2 \times 2(\nu + 2)} = \frac{a_0}{2^4 \times 1 \times 2(\nu + 1)(\nu + 2)}.$$

Or, in general,

$$a_{2m} = (-1)^m \frac{a_0}{2^{2m} m! (\nu + 1)(\nu + 2) \dots (\nu + m)}.$$

We now substitute the values of the coefficients into (18.85):

$$y_1(x) = a_0 x^\nu \left(1 + \sum_{m=1}^{\infty} (-1)^m \frac{x^{2m}}{2^{2m} m! (\nu + 1)(\nu + 2) \dots (\nu + m)} \right). \tag{18.87}$$

It can easily be verified that the series on the right of (18.87) converges in any case on the positive x-axis and defines there the function $y_1(x)$, i.e., a particular solution of the Bessel equation.

Consider now the case when $\sigma = -\nu$. If ν is no positive integer, then we can write the second particular solution that is deduced from (18.87) by the change of ν by $-\nu$ (in (18.84) ν appears evenly):

$$y_2(x) = a_0 x^{-\nu} \left(1 + \sum_{m=1}^{\infty} (-1)^m \frac{x^{2m}}{2^{2m} m! (-\nu + 1)(-\nu + 2) \dots (-\nu + m)} \right). \tag{18.87$'$}$$

(If ν is a positive integer, then the solution (18.87′) is no longer valid, since beginning with a certain number one of the factors in the denominator in (18.87′) will be zero.) The series on the right of (18.87′) also converges for all values of $x > 0$. The solutions $y_1(x)$ and $y_2(x)$ are linearly independent. Really, their ratio

$$\frac{y_2(x)}{y_1(x)} = x^{-2\nu} \frac{1 - \dfrac{x^2}{2^2(-\nu + 1)} + \dots}{1 - \dfrac{x^2}{2^2(\nu + 1)} + \dots}$$

is not constant.

For our further discussion we will need some of the properties of Euler's Γ-function. The latter is defined as follows:

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} \, dx, \quad \operatorname{Re} p > 0.$$

Integrating by parts we obtain the basic functional equation for the Γ-function

$$\Gamma(p + 1) = p\Gamma(p). \tag{18.88}$$

Since $\Gamma(1) = 1$, then $\Gamma(2) = 1 \cdot \Gamma(1) = 1$, $\Gamma(3) = 2\Gamma(2) = 2!$, and in general

$$\Gamma(n + 1) = n! \quad (n = 0, 1, 2, \dots).$$

It can be shown that $\Gamma(1/2) = \sqrt{\pi}$. Using the functional equation (18.88) we can derive the gamma-function for negative values of the argument. If we represent (18.88) in the form $\Gamma(p) = \Gamma(p+1)/p$, we notice that for small p we have $\Gamma(p) \approx 1/p$.

Similarly, if m is a positive integer, then for p close to $-m$ we have

$$\Gamma(p) \approx \frac{(-1)^m}{m!}\frac{1}{p+m}.$$

It can be shown that $\Gamma(p) \neq 0$ for all p, therefore the function $1/\Gamma(p)$ will be continuous for all p if we put

$$1/\Gamma(-m) = 0 \quad (m = 0, 1, 2, \dots).$$

Let us return to the solution of the Bessel equation (18.84). Here a_0 is still arbitrary. If $\nu \neq -n$, where $n > 0$ is an integer, then putting

$$a_0 = \frac{1}{2^\nu \Gamma(\nu+1)},$$

gives

$$a_{2m} = (-1)^m \frac{1}{2^{2m+\nu} m!(\nu+1)(\nu+2)\dots(\nu+m)\Gamma(\nu+1)}$$
$$= (-1)^m \frac{1}{2^{2m+\nu}\Gamma(m+1)\Gamma(\nu+m+1)}.$$

Substituting this into (18.86) gives

$$y_1(x) = \sum_{m=0}^{\infty} (-1)^m \frac{1}{\Gamma(m+1)\Gamma(\nu+m+1)}\left(\frac{x}{2}\right)^{2m+\nu}. \tag{18.89}$$

Series (18.89) defines the function

$$J_\nu(x) = \sum_{m=0}^{\infty} (-1)^m \frac{1}{\Gamma(m+1)\Gamma(\nu+m+1)}\left(\frac{x}{2}\right)^{2m+\nu}, \tag{18.90}$$

which is a solution of the Bessel equation and is called the *Bessel function* of the first kind of order ν.

The series

$$J_{-\nu}(x) = \sum_{m=0}^{\infty} (-1)^m \frac{1}{\Gamma(m+1)\Gamma(-\nu+m+1)}\left(\frac{x}{2}\right)^{2m-\nu}$$

corresponds to the case of $\sigma = -\nu$ (ν is a noninteger) and defines the second solution of (18.84), which is linearly independent of $J_\nu(x)$.

We thus conclude that if ν is a noninteger ($\nu \neq 0, \pm 1, \pm 2, \ldots$), then the functions $J_\nu(x)$ and $J_{-\nu}(x)$ form a fundamental set of solutions of the Bessel equation (18.84) and its general solution will then have the form

$$y = C_1 J_\nu(x) + C_2 J_{-\nu}(x).$$

For integral ν, we have the linear dependence

$$J_{-n}(x) = (-1)^n J_n(x). \tag{18.91}$$

Indeed, we have

$$J_{-n}(x) = \sum_{m=0}^{\infty} (-1)^m \frac{1}{\Gamma(m+1)\Gamma(m-n+1)} \left(\frac{x}{2}\right)^{2m-n}.$$

The first n terms of the series vanish, since $1/\Gamma(m-n+1) = 0$ at $m = 0, 1, \ldots, n-1$, and $1/\Gamma(n-n+1) = 1$. Introducing the notation $m = k+n$, we find

$$J_{-n}(x) = \sum_{k=0}^{\infty} (-1)^{k+n} \frac{1}{\Gamma(k+n+1)\Gamma(k+1)} \left(\frac{x}{2}\right)^{2k+n}$$

$$= (-1)^n \sum_{k=0}^{\infty} (-1)^k \frac{1}{\Gamma(k+1)\Gamma(k+n+1)} \left(\frac{x}{2}\right)^{2k+n} = (-1)^n J_n(x).$$

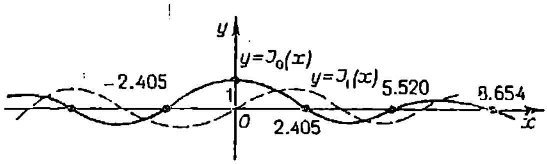


Fig. 18.4

We write the series for the Bessel function of the first kind of order zero ($n = 0$) and order one ($n = 1$):

$$J_0(x) = 1 - \left(\frac{x}{2}\right)^2 + \frac{1}{(2!)^2}\left(\frac{x}{2}\right)^4 - \frac{1}{(3!)^2}\left(\frac{x}{2}\right)^6 + \ldots,$$

$$J_1(x) = \frac{x}{2} - \frac{1}{2!}\left(\frac{x}{2}\right)^3 + \frac{1}{2!3!}\left(\frac{x}{2}\right)^5 - \ldots.$$

The functions $J_0(x)$ and $J_1(x)$ (Fig. 18.4) often occur in applications. For them detailed tables are available.

Recurrence formulas. Using the formula (18.90) we make sure, by direct check, that

$$\frac{d}{dx}(x^{\nu}J_{\nu}(x)) = x^{\nu}J_{\nu-1}(x). \tag{18.92}$$

In exactly the same manner we find

$$\frac{d}{dx}(x^{-\nu}J_{\nu}(x)) = -x^{-\nu}J_{\nu+1}(x). \tag{18.93}$$

Expanding the derivatives in the left-hand sides of (18.92) and (18.93), we obtain

$$J_{\nu}'(x) + \frac{\nu}{x}J_{\nu}(x) = J_{\nu-1}(x), \tag{18.94}$$

$$J_{\nu}'(x) - \frac{\nu}{x}J_{\nu}(x) = -J_{\nu+1}(x). \tag{18.95}$$

By adding and subtracting (18.94) and (18.95), we obtain two important *recurrence formulas*:

$$J_{\nu}'(x) = \frac{1}{2}[J_{\nu-1}(x) - J_{\nu+1}(x)], \tag{18.96}$$

$$J_{\nu+1}(x) + J_{\nu-1}(x) = \frac{2\nu}{x}J_{\nu}(x). \tag{18.97}$$

Formula (18.96) indicates that derivatives of Bessel functions are expressed through Bessel functions. It follows from (18.97) that, knowing $J_{\nu}(x)$ and $J_{\nu-1}(x)$, we can find $J_{\nu+1}(x)$. Specifically, all Bessel functions of whole numbers are expressed through $J_0(x)$ and $J_1(x)$. Relation (18.91) comes in handy here. At $\nu = 1$ we find from (18.97), e.g.,

$$J_2(x) = \frac{2}{x}J_1(x) - J_0(x).$$

Bessel functions for half-integer *n*. Consider a special class of Bessel functions with half-integral odd n. This class occurs in applications and is noted for the fact that in the case under discussion Bessel functions can be expressed through elementary functions. So at $\nu = 1/2$ we readily find

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\sin x.$$

Likewise, at $\nu = -1/2$

$$J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\cos x.$$

The above furmulas can be rewritten as

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}}\cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right), \quad \nu = \pm\frac{1}{2}. \tag{18.98}$$

Using the recurrence formula (18.97) we find, for example, that

$$J_{3/2}(x) = \frac{1}{x}J_{1/2}(x) - J_{-1/2}(x) = \sqrt{\frac{2}{\pi x}}\left(\frac{\sin x}{x} - \cos x\right),$$

and so on.

Zeros of Bessel functions. In many applications it would be instructive to have an idea of the distribution of *zeros* of Bessel functions. The zeros of $J_{1/2}(x)$ and $J_{-1/2}(x)$ coincide with the zeros of $\sin x$ and $\cos x$, respectively. It can readily be shown that for large x we have the asymptotic representation (cf. (18.98))

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}}\cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) + O(x^{-3/2}), \quad x \to +\infty, \tag{18.99}$$

that holds for any integer or noninteger ν. (Here $f(x) = O(\varphi(x))$ means that the ratio $f(x)/\varphi(x)$ remains bounded as $x \to \infty$.) Formula (18.99) shows the behaviour of the Bessel function with increasing argument. This oscillating function becomes zero an infinite number of times and the amplitude of the oscillation tends to zero as $x \to +\infty$.

The distribution of zeros of the Bessel function with positive integral n, i.e., the roots of the equation

$$J_n(x) = 0 \quad (n = 0, 1, 2, \ldots)$$

is established by the following theorem.

Theorem 18.18. *The function $J_n(x)$ $(n = 0, 1, 2, \ldots)$ has no complex zeros, but has an infinite number of real zeros, arranged symmetrically about the point $x = 0$, which is one of them if $n = 1, 2, \ldots$. All the zeros of the function are simple, except for $x = 0$, which at $n = 1, 2, \ldots$ is a zero of multiplicity n.*

Orthogonality and norm of Bessel functions. We first consider the property of *orthogonality*. It can readily be verified that the equation

$$x^2y'' + xy' + (\lambda^2x^2 - \nu^2)y = 0, \tag{18.100}$$

where λ is some nonzero numerical parameter, is satisfied by the Bessel function $J_\nu(\lambda x)$.

We rewrite (18.100) in the form

$$y'' + \frac{1}{x}y' + \left(\lambda^2 - \frac{\nu^2}{x^2}\right)y = 0 \tag{18.101}$$

and denote $y_1 = J_\nu(\lambda_1 x)$, $y_2 = J_\nu(\lambda_2 x)$, where λ_1, λ_2 are some values of λ.

We will then have the identities

$$y_1'' + \frac{1}{x} y_1' + \left(\lambda_1^2 - \frac{\nu^2}{x^2}\right) y_1 \equiv 0,$$

$$y_2'' + \frac{1}{x} y_2' + \left(\lambda_2^2 - \frac{\nu^2}{x^2}\right) y_2 \equiv 0.$$

Multiplying the first one by $y_2(x)$, the second by $y_1(x)$ and subtracting one from the other, we will get

$$y_1'' y_2 - y_1 y_2'' + \frac{1}{x} (y_1' y_2 - y_1 y_2') + (\lambda_1^2 - \lambda_2^2) y_1 y_2 \equiv 0.$$

Multiplying this identity by x, we notice that it can be written as

$$\frac{d}{dx} [x(y_1' y_2 - y_1 y_2')] \equiv (\lambda_2^2 - \lambda_1^2) x y_1 y_2.$$

Integrating this with respect to x from 0 to 1, we will have

$$[x(y_1' y_2 - y_1 y_2')]\Big|_{x=0}^{x=1} = (\lambda_2^2 - \lambda_1^2) \int_0^1 x y_1(x) y_2(x)\, dx.$$

or

$$\lambda_1 J_\nu'(\lambda_1) J_\nu(\lambda_2) - \lambda_2 J_\nu'(\lambda_2) J_\nu(\lambda_1)$$
$$= (\lambda_2^2 - \lambda_1^2) \int_0^1 x J_\nu(\lambda_1 x) J_\nu(\lambda_2 x)\, dx. \qquad (18.102)$$

(1) Let $\lambda_1 \neq \lambda_2$. Then from (18.102) it follows that if λ_1, λ_2 are zeros of $J_\nu(x)$, then the left-hand side of (18.102) and, hence, the right-hand side as well, are zero. Then

$$\int_0^1 x J_\nu(\lambda_1 x) J_\nu(\lambda_2 x)\, dx = 0.$$

This implies that, by definition, $J_\nu(\lambda_1 x)$ and $J_\nu(\lambda_2 x)$ are orthogonal with weight $\varrho(x) = x$ on the interval [0, 1].

The Bessel function $J_\nu(x)$ has the countable set of zeros

$$0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots$$

and hence, the set of functions

$$J_\nu(\lambda_1 x),\ J_\nu(\lambda_2 x),\ \ldots,\ J_\nu(\lambda_n x),\ \ldots, \qquad (18.103)$$

where ν is fixed, is a system orthogonal on the interval [0, 1] with weight $\varrho(x) = x$:

$$\int_0^1 x J_\nu(\lambda_i x) J_\nu(\lambda_j x)\, dx = 0, \quad i \neq j. \qquad (18.104)$$

(2) If λ_1 and λ_2 are the roots of the equation

$$J_\nu'(x) = 0,$$

then at $\lambda_1 \neq \lambda_2$ we from (18.102) also have

$$\int_0^1 x\, J_\nu(\lambda_1 x)\, J_\nu(\lambda_2 x)\, dx = 0.$$

Accordingly, the set of functions $\{J_\nu(\lambda_n x)\}_{n=1}^{\infty}$, where λ_n are roots of the equation $J_\nu'(x) = 0$, is orthogonal on the interval $[0, 1]$ with weight $\varrho(x) = x$.

(3) Let λ_1 and λ_2 be the roots of the equation

$$\frac{x J_\nu'(x)}{J_\nu(x)} = h, \tag{18.105}$$

where h is some fixed number.

Equation (18.105) is used in mathematical physics and for $\nu > -1$ it has an infinite number of positive roots, but it has no complex roots, save for the case of $(-h + \nu) < 0$, where there are two purely imaginary roots.

If we write the left-hand side of (18.102) in the form

$$\lambda_1 J_\nu'(\lambda_1) J_\nu(\lambda_2) - \lambda_2 J_\nu'(\lambda_2) J_\nu(\lambda_1) = J_\nu(\lambda_1) J_\nu(\lambda_2) \left[\frac{\lambda_1 J_\nu'(\lambda_1)}{J_\nu(\lambda_1)} - \frac{\lambda_2 J_\nu'(\lambda_2)}{J_\nu(\lambda_2)}\right],$$

we will see that Bessel functions are orthogonal in zeros of the linear combination $xJ_\nu'(x) - hJ_\nu(x) = 0$ of the Bessel function and its derivative

$$\int_0^1 xJ_\nu(\lambda_i x) J_\nu(\lambda_j x) dx = 0, \qquad i \neq j,$$

where λ_k $(k = 1, 2, \ldots)$ are the roots of (18.105).

The quantity

$$\| J_\nu(\lambda_x) \| = \left(\int_0^1 x J_\nu^2(\lambda x) dx\right)^{1/2}$$

is called the *norm* of the Bessel function $J_\nu(\lambda x)$.

Using the equality (18.102), we can show that

$$\| J_\nu(\lambda x) \|^2 = \frac{1}{2}\left[J_\nu'^2(\lambda) + \left(1 - \frac{\nu^2}{\lambda^2}\right) J_\nu^2(\lambda)\right]. \tag{18.106}$$

In particular, for $J_0(\lambda x)$, where λ is a zero of the Bessel function we have

$$\| J_0(\lambda x) \|^2 = \frac{1}{2} J_0'^2(\lambda) = \frac{1}{2} J_1^2(\lambda),$$

since $J_0'(\lambda) = -J_1(\lambda)$.

Neumann (Weber) functions. Any nontrivial solution of the Bessel equation (18.84) is called a *cylindrical function*. When ν is not an integer, the functions $J_\nu(x)$ and $J_{-\nu}(x)$ form a fundamental set of solutions of the Bessel equation (18.84). When $\nu = n$, i.e., n is an integer we have the linear dependence

$$J_{-n}(x) = (-1)^n J_n(x).$$

To suppelement the solution $J_n(x)$ by one that is not proportional to it, we proceed as follows. When ν is not an integer, we form the function

$$N_\nu(x) = \frac{J_\nu(x)\cos \pi\nu - J_{-\nu}(x)}{\sin \pi\nu}. \tag{18.107}$$

It is a linear combination of solutions of the linear homogeneous equation (18.84), and so it is itself a solution of the equation.

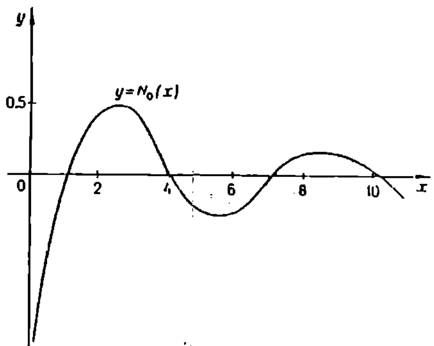


Fig. 18.5

In the limit as $\nu \to n$, we, by L'Hospital's rule, will have

$$N_n(x) = \left.\frac{(-1)^n \dfrac{\partial J_\nu}{\partial \nu} - \dfrac{\partial J_{-\nu}}{\partial \nu}}{(-1)^n \pi}\right|_{\nu = n}.$$

One distinction of $N_n(x)$ (Bessel function of the second kind) is the presence of a singularity at the origin of coordinates (Fig. 18.5).

$$N_0(x) \sim \frac{2}{\pi}\ln\frac{2}{\gamma x}, \qquad \gamma = 1.781...$$

$$x \to 0 + 0.$$

$$N_n(x) \sim -\frac{(n-1)!}{\pi}\left(\frac{2}{x}\right)^n \qquad (n = 1, 2, ...)$$

The solution $N_n(x)$ of the Bessel equation (18.84) at $\nu = n$ constitutes

together with $J_n(x)$ a fundamental set of solutions of the equation

$$x^2y'' + xy' + (x^2 - n^2)y = 0.$$

The function $N_n(x)$ is also called the *Neumann* (or *Weber*) *function*.

At sufficiently large x

$$\left.\begin{aligned} J_\nu(x) &\sim \sqrt{\frac{2}{\pi x}}\cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \\ N_\nu(x) &\sim \sqrt{\frac{2}{\pi x}}\sin\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \end{aligned}\right\} \quad x \to +\infty.$$

Table 18.2 Forms of particular solutions of inhomogeneous linear equations with constant coefficients for various right-hand sides

Right-hand side* of equations	Roots of characteristic equations	Forms of particular solution
1. $P_m(x)$	Number 0 is no root of characteristic equation	$\tilde{P}_m(x)$
	Number 0 is a root of characteristic equation of multiplicity $r \geqslant 1$	$x^r\tilde{P}_m(x)$
2. $e^{ax}P_m(x)$	Number a is no root of characteristic equation	$e^{ax}\tilde{P}_m(x)$
	Number a is a root of characteristic equation of multiplicity $r \geqslant 1$	$x^re^{ax}\tilde{P}_m(x)$
3. $P_m(x)\cos\beta x + Q_s(x)\sin\beta x$	Numbers $\pm i\beta$ are no roots of characteristic equation	$\tilde{P}_k(x)\cos\beta x + \tilde{Q}_k(x)\sin\beta x$, $k = \max\{m, s\}$
	Numbers $\pm i\beta$ are roots of characteristic equation of multiplicity r	$x^r(\tilde{P}_k(x)\cos\beta x + \tilde{Q}_k(x)\sin\beta x)$
4. $e^{\alpha x}[P_m(x)\cos\beta x + Q_s(x)\sin\beta x]$	Numbers $\alpha \pm i\beta$ are no roots of characteristic equation	$e^{\alpha x}(\tilde{P}_k(x)\cos\beta x + \tilde{Q}_k(x)\sin\beta x)$, $k = \max\{m, s\}$
	Numbers $\alpha \pm i\beta$ are roots of characteristic equation of multiplicity r	$x^re^{\alpha x}(\tilde{P}_k(x)\cos\beta x + \tilde{Q}_k(x)\sin\beta x)$

* The first three kinds of right-hand sides are special cases of the fourth.

This suggests that at large distances from the origin of coordinates the cylindrical functions of the first and second kinds are related as those of sine and cosine. Owing to the factor $1/\sqrt{x}$ the functions decay as x grows. These functions are convenient to represent standing cylindrical waves.

In analogy with exponential functions (Euler formulas) we can construct a linear combination of $J_\nu(x)$ and $N_\nu(x)$ to obtain functions associated with running waves. This brings us to the Bessel functions of the third kind or *Hankel functions* given by

$$H_\nu^{(1)}(x) = J_\nu(x) + iN_\nu(x),$$
$$H_\nu^{(2)}(x) = J_\nu(x) - iN_\nu(x).$$

Exercises

Find the general solution of the equations:

1. $(1 + x^2)y'' + 2xy' = x^3$. 2. $y^{(4)}\tanh x = y'''$. 3. $yy'' = y'^2$. 4. $yy'' + y'^2 = 1$.

Find the solution of the Cauchy problem:

5. $y'' + 18\sin y\cos^3 y = 0$, $y(0) = 0$, $y'(0) = 3$.
6. $y'' = 18y^3$, $y(1) = 1$, $y'(1) = 3$.
7. $y^3y'' = 4(y^4 - 1)$, $y(0) = \sqrt{2}$, $y'(0) = \sqrt{2}$.

Integrate the equations. Where required, find particular solutions.

8. $y'' - 4y' + 4y = 0$. 9. $y'' - 2y' - 3y = 0$. 10. $y'' + 2y' + 5y = 0$. 11. $y'' - 3y' + 2y = 0$, $y(0) = 0$, $y'(0) = 1$. 12. $y''' - y' = 0$. 13. $y''' - y = 0$, $y(1) = y'(1) = y''(1) = 0$. 14. $y^{(4)} - y = 0$. 15. $y^{(4)} + y = 0$. 16. $y^{(5)} = 0$.

Find the general solutions for

17. $y'' + y = 1$. 18. $y'' - 2y' + y = x + 1$. 19. $y'' - y = 4xe^x$. 20. $y'' - 2y' + y = 2e^x$. 21. $y'' + y' + y = \sin x$. 22. $y'' + y = \cos x - 2\sin x$. 23. $y'' + 4y = 3x\sin x$. 24. $y'' + 4y = 2\sin^2 x$.

Find the form of particular solutions for

25. $y''' - y' = 3 + xe^x + x^2\sin x$. 26. $y''' + y' = 2 + x + x^2e^x + x\sin x$. 27. $y''' - y'' = 1 + xe^x + 2x\cos x$. 28. $y''' + y'' = x + xe^x + x\sin x$. 29. $y'' - y = 1 + xe^x + e^x\cos x$.

Integrate by variation of constants

30. $y'' + y = 1/\cos x$. 31. $y'' - y' = e^{2x}\sin e^x$.

Integrate the following Euler equations:

32. $x^2y'' - xy' - 3y = 0$. 33. $x^2y'' + xy' + y = 0$.

Answers

1. $y = \dfrac{x^3}{12} - \dfrac{x}{4} + C_1\tan^{-1}x + C_2$. 2. $y = C_1\cosh x + C_2x^2 + C_3x + C_4$. 3. $y = C_2e^{C_1x}$. 4. $y^2 = (x + C_2)^2 + C_1$. 5. $y = \tan^{-1}3x$. 6. $y = \dfrac{1}{4 - 3x}$.

7. $y = \sqrt{1 + e^{4x}}$. 8. $y = C_1e^{2x} + C_2xe^{2x}$. 9. $y = C_1e^{-x} + C_2e^{3x}$. 10. $y = C_1e^{-x}\cos 2x + C_2e^{-x}\sin 2x$. 11. $y = e^{2x} - e^{x}$. 12. $y = C_1 + C_2e^{-x} + C_3e^{x}$. 13. $y \equiv 0$. 14. $y = C_1e^{x} + C_2e^{-x} + C_3\cos x + C_4\sin x$. 15. $y = C_1e^{x\sqrt{2}/2}\cos\frac{\sqrt{2}}{2}x + C_2e^{x\sqrt{2}/2}\sin\frac{\sqrt{2}}{2}x + C_3e^{-x\sqrt{2}/2}\cos\frac{\sqrt{2}}{2}x + C_4e^{-x\sqrt{2}/2}\sin\frac{\sqrt{2}}{2}x$. 16. $y = C_1 + C_2x + C_3x^2 + C_4x^3 + C_5x^4$. 17. $y = C_1\cos x + C_2\sin x + 1$. 18. $y = C_1e^{x} + C_2xe^{x} + x + 3$. 19. $y = C_1e^{x} + C_2e^{-x} + (x^2 - x)e^{x}$. 20. $y = C_1e^{x} + C_2xe^{x} + x^2e^{x}$. 21. $y = C_1e^{-x/2}\cos\frac{\sqrt{3}}{2}x + C_2e^{-x/2}\sin\frac{\sqrt{3}}{2}x - \cos x$. 22. $y = C_1\cos x + C_2\sin x + x\left(\cos x + \frac{1}{2}\sin x\right)$. 23. $y = C_1\cos 2x + C_2\sin 2x + x\sin x - \frac{2}{3}\cos x$. 24. $y = C_1\cos 2x + C_2\sin 2x + \frac{1}{4} - \frac{x}{4}\sin 2x$. 25. $y = Ax + x(A_1x + B_1)e^{x} + (A_2x^2 + B_2x + D_2)\cos x + (A_3x^2 + B_3x + D_3)\sin x$. 26. $y = x(Ax + B) + (A_1x^2 + B_1x + D_1)e^{x} + x[(A_2x + B_2)\cos x + (A_3x + B_3)\sin x]$. 27. $y = Ax^2 + x(A_1x + B_1)e^{x} + (A_2x + B_2)\cos x + (A_3x + B_3)\sin x$. 28. $y = (Ax + B)x^2 + (A_1x + B_1)e^{x} + (A_2x + B_2)\cos x + (A_3x + B_3)\sin x$. 29. $y = A + (A_1x + B_1)e^{2x} + e^{x}(A_2\cos x + B_2\sin x)$. 30. $y = C_1\cos x + C_2\sin x + \cos x\ln|\cos x| + x\sin x$. 31. $y = C_1 + C_2e^{x} - \sin e^{x}$. 32. $y = C_1x^3 + \frac{C_2}{x}$. 33. $y = C_1\cos(\ln x) + C_2\sin(\ln x)$, $x > 0$.

Chapter 19

Systems of Differential Equations

19.1 Essentials. Definitions

A system of differential equations models even the simplest problem of particle dynamics: given the forces acting on a particle, find the law of motion, i.e., find the functions $x = x(t)$, $y = y(t)$, $z = z(t)$, which express the dependence of the coordinates of a moving particle on time. The system that results in the general case has the form

$$\begin{cases} \dfrac{d^2x}{dt^2} = f\left(t,\ x,\ y,\ z,\ \dfrac{dx}{dt},\ \dfrac{dy}{dt},\ \dfrac{dz}{dt}\right), \\[2ex] \dfrac{d^2y}{dt^2} = g\left(t,\ x,\ y,\ z,\ \dfrac{dx}{dt},\ \dfrac{dy}{dt},\ \dfrac{dz}{dt}\right), \\[2ex] \dfrac{d^2z}{dt^2} = h\left(t,\ x,\ y,\ z,\ \dfrac{dx}{dt},\ \dfrac{dy}{dt},\ \dfrac{dz}{dt}\right). \end{cases} \tag{19.1}$$

Here x, y, z are the coordinates of a travelling particle, t is time; f, g, h are known functions of respective arguments.

A system of the type (19.1) is known as a *canonic system*. Turning to the general case of the system of m differential equations with m unknown functions $x_1(t)$, $x_2(t)$, $\ldots$, $x_m(t)$ of t, we will treat as canonic the system of the form

$$x_i^{(k_i)} = f_i(t,\ x_1,\ x_1',\ \ldots,\ x_1^{(k_1-1)},\ \ldots,\ x_m,\ x_m',\ \ldots,\ x_m^{(k_m-1)}),$$
$$i = 1,\ 2,\ \ldots,\ m \tag{19.2}$$

which is solvable for the highest-order derivatives. A system of first-order equations solvable for the derivatives of the desired functions

$$x_i' = f_i(t,\ x_1,\ x_2,\ \ldots,\ x_n) \qquad (i = 1,\ 2,\ \ldots,\ n) \tag{19.3}$$

is called a *normal system*

If x_i', x_i'', $\ldots$, $x_i^{(k_i-1)}$ in (19.2) are taken to be new auxiliary functions, then the general canonic system (19.2) can be replaced by an equivalent

normal system consisting of $N = k_1 + k_2 + \ldots + k_m$ equations. Therefore, it is sufficient to consider only normal systems.

For example, one equation $d^2x/dt^2 = -x$ is a special case of the canonic system. Setting $dx/dt = y$, we will have $dy/dt = -x$ from the original equation. As a result, we will have the normal system of equations

$$\begin{cases} \dfrac{dx}{dt} = y, \\ \dfrac{dy}{dt} = -x \end{cases}$$

equivalent to the original equation.

Definition. Any system of n functions

$$x_1 = x_1(t),\ x_2 = x_2(t),\ \ldots,\ x_n = x_n(t), \tag{19.4}$$

differentiable on the interval $a < t < b$, such that it turns the equations of (19.3) into identities in t on the interval (a, b) is called a *solution of the normal system* (19.3) for t defined on the interval (a, b).

The Cauchy problem for the system (19.3) is formulated as follows: find the solution (19.4) of the system such that at $t = t_0$ it obeys the initial conditions

$$x_1\big|_{t=t_0} = x_1^0, \qquad x_2\big|_{t=t_0} = x_2^0,\ \ldots,\ x_n\big|_{t=t_0} = x_n^0. \tag{19.5}$$

Theorem 19.1 (on existence and uniqueness of the solution of the Cauchy problem). *Let* (19.3) *be a normal system of differential equations. Suppose that the functions* $f_i\,(t, x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, n$, *are defined in a certain* $(n+1)$*-dimensional domain* D *of the variables* $t, x_1, x_2, \ldots, x_n$. *If there exists a neighbourhood* Ω *of a point* $M_0\,(t_0, x_1^0, x_2^0, \ldots, x_n^0)$ *where* f_i *are continuous in the multitude of arguments and have bounded partial derivatives in* $x_1, x_2, \ldots, x_n$, *then there will be an interval* $t_0 - h_0 < t < t_0 + h_0$ *where there exists a unique solution of the normal system* (19.3) *satisfying the initial conditions* $x_1\big|_{t=t_0} = x_1^0$, $x_2\big|_{t=t_0} = x_2^0, \ldots, x_n\big|_{t=t_0} = x_n^0$.

Definition. A system of n functions

$$x_i = x_i\,(t, C_1, C_2, \ldots, C_n) \qquad (i = 1, 2, \ldots, n) \tag{19.6}$$

of t and n arbitrary constants $C_1, C_2, \ldots, C_n$, is called the *general solution of the normal system* (19.3) in a certain domain Ω where there exists a unique solution of the Cauchy problem, if

(1) for any permissible values of $C_1, C_2, \ldots, C_n$ the system (19.6) turns equations (19.3) into identities;

(2) in Ω the functions (19.6) solve any Cauchy problem.

Solutions that are deduced from the general one for concrete values of $C_1, C_2, \ldots, C_n$ are called *particular solutions*.

We will turn for definiteness to the normal system of two equations

$$\begin{cases} \dfrac{dx_1}{dt} = f_1(t, x_1, x_2), \\ \dfrac{dx_2}{dt} = f_2(t, x_1, x_2). \end{cases} \tag{19.7}$$

We will treat the system of values of t, x_1, x_2 as the Cartesian coordinates of a point in a three-dimensional space with the system of coordinates Otx_1x_2. The solution $x_1 = x_1(t)$, $x_2 = x_2(t)$ of the system (19.7), which at $t = t_0$ assumes the values x_1^0, x_2^0, defines in the space a line passing through the point $M_0(t_0, x_1^0, x_2^0)$. This line is known as the *integral curve* of the normal system (19.7). The Cauchy problem for (19.7) can be given the following geometrical treatment: in the space of t, x_1, x_2 it is required to find the integral curve passing through a given point $M_0(t_0, x_1^0, x_2^0)$ (Fig. 19.1). Theorem 19.1 establishes the existence and uniquenes of such a curve.

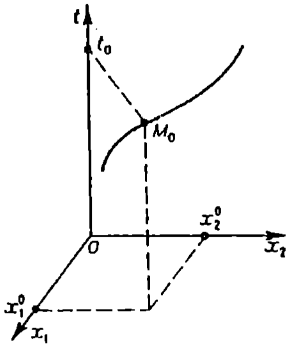


Fig. 19.1

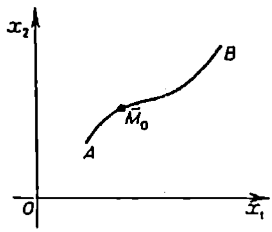


Fig. 19.2

The normal system (19.7) and its solution can also be treated as follows: we will view the independent variable t as a parameter, and the solution $x_1 = x_1(t)$, $x_2 = x_2(t)$ of the system as parametric equations of the curve in the x_1x_2-plane. This plane of the variables x_1, x_2 is called the *phase plane*. In the phase plane the solution $x_1 = x_1(t)$, $x_2 = x_2(t)$ of (19.7), which at $t = t_0$ takes on the initial values x_1^0, x_0^2, is shown by the curve AB, passing through $\tilde{M}_0(x_1^0, x_2^0)$ (Fig. 19.2). This curve is termed the *path* (*trajectory*) *of the system* (*phase path*). The path of (19.7) is the projection of the integral curve on the phase plane. From an integral curve the phase path can be determined uniquely, but not vice versa.

19.2 Methods of Integration of Systems of Differential Equations

One of the integration methods is *integration by elimination*. A special case of a canonic system is one equation of order n solvable for the highest-order derivative

$$x^{(n)} = f(t, x, x', \ldots, x^{(n-1)}).$$

Introducing the new variables $x_1 = x'(t)$, $x_2 = x''(t)$, $\ldots$, $x_{n-1} = x^{(n-1)}(t)$, we replace this equation by the following normal system of n equations

$$\begin{cases} \dfrac{dx}{dt} = x_1, \\ \dfrac{dx_1}{dt} = x_2, \\ \vdots \\ \dfrac{dx_{n-2}}{dt} = x_{n-1}, \\ \dfrac{dx_{n-1}}{dt} = f(t, x, x_1, \ldots, x_{n-1}), \end{cases} \tag{19.8}$$

i.e., one equation of nth order is equivalent to the normal system (19.8).

The reverse is also true, i.e., generally speaking, a normal system of n equations of the first order is equivalent to one equation of order n. It is on this that the elimination method of integration of a system of differential equations is based. It is done as follows. Let there be a normal system of differential equations

$$\begin{cases} \dfrac{dx_1}{dt} = f_1(t, x_1, x_2, \ldots, x_n), \\ \dfrac{dx_2}{dt} = f_2(t, x_1, x_2, \ldots, x_n), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{dx_n}{dt} = f_n(t, x_1, x_2, \ldots, x_n). \end{cases} \tag{19.9}$$

We differentiate, for example, the first equation in (19.9) with respect to t:

$$\frac{d^2x}{dt^2} = \frac{\partial f_1}{\partial t} + \frac{\partial f_1}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f_1}{\partial x_2}\frac{dx_2}{dt} + \ldots + \frac{\partial f_1}{\partial x_n}\frac{dx_n}{dt}.$$

Substituting dx_i/dt on the right-hand side by the expressions $f_i(t, x_1, x_2, \ldots, x_n)$ yields

$$\frac{d^2x_1}{dt^2} = \frac{\partial f_1}{\partial t} + \frac{\partial f_1}{\partial x_1} f_1 + \frac{\partial f_1}{\partial x_2} f_2 + \ldots + \frac{\partial f_1}{\partial x_n} f_n(t, x_1, x_2, \ldots, x_n),$$

i.e., an expression of the form

$$\frac{d^2x_1}{dt^2} = F_2(t, x_1, x_2, \ldots, x_n). \tag{19.10}$$

This equation is again differentiated with respect to t. Using (19.9), we get

$$\frac{d^3x_1}{dt^3} = \frac{\partial F_2}{\partial t} + \frac{\partial F_2}{\partial x_1} f_1 + \ldots + \frac{\partial F_2}{\partial x_n} f_n,$$

or

$$\frac{d^3x_1}{dt^3} = F_3(t, x_1, x_2, \ldots, x_n).$$

If this process is continued, we get

$$\frac{d^4x_1}{dt^4} = F_4(t, x_1, x_2, \ldots, x_n),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{d^{n-1}x_1}{dt^{n-1}} = F_{n-1}(t, x_1, x_2, \ldots, x_n),$$

$$\frac{d^nx_1}{dt^n} = F_n(t, x_1, x_2, \ldots, x_n).$$

We assume that the determinant (the Jacobian of the system of functions $f_1, F_2, \ldots, F_{n-1}$)

$$\frac{D(f_1, F_2, \ldots, F_{n-1})}{D(x_2, x_3, \ldots, x_n)} = \begin{vmatrix} \dfrac{\partial f_1}{\partial x_2} & \dfrac{\partial f_1}{\partial x_3} & & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial F_2}{\partial x_2} & \dfrac{\partial F_2}{\partial x_3} & & \dfrac{\partial F_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial F_{n-1}}{\partial x_2} & \dfrac{\partial F_{n-1}}{\partial x_3} & & \dfrac{\partial F_{n-1}}{\partial x_n} \end{vmatrix}$$

is nonzero for the values of $x_2, x_3, \ldots, x_n$ in question, viz.,

$$\frac{D(f_1, F_2, \ldots, F_{n-1})}{D(x_2, x_3, \ldots, x_n)} \neq 0. \tag{19.11}$$

The system of equations containing the first equation in (19.9) and the equations

$$\frac{d^2x_1}{dt^2} = F_2, \ \ldots, \ \frac{d^{n-1}x_1}{dt^{n-1}} = F_{n-1}(t, x_1, x_2, \ldots, x_n)$$

will be solvable for $x_2, x_3, \ldots, x_n$, which are expressed in terms of $t, x_1,$

$dx_1/dt, \ldots, d^{n-1}x_1/dt^{n-1}$. Entering these expressions into

$$\frac{d^n x_1}{dt^n} = F_n(t, x_1, x_2, \ldots, x_n)$$

yields one equation of order n

$$\frac{d^n x_1}{dt^n} = \Phi\left(t, x_1, \frac{dx_1}{dt}, \ldots, \frac{d^{n-1}x_1}{dt^{n-1}}\right). \tag{19.12}$$

It follows from the way it was derived that if $x_1(t), x_2(t), \ldots, x_n(t)$ are solutions of (19.9), then $x_1(t)$ will be a solution of (19.12). Conversely, let $x_1(t)$ be a solution to (19.12). If we differentiate this solution with respect to t, we get $\frac{dx_1}{dt}, \ldots, \frac{d^{n-1}x_1}{dt^{n-1}}$. We get these values as known functions of t in the system of equations

$$\frac{dx_1}{dt} = f_1(t, x_1, x_2, \ldots, x_n), \frac{d^2x_1}{dt^2} = F_2(t, x_1, x_2, \ldots, x_n), \ldots,$$
$$\frac{d^{n-1}x_1}{dt^{n-1}} = F_{n-1}(t, x_1, x_2, \ldots, x_n).$$

By our assumption, this system can be solved for $x_2, x_3, \ldots, x_n$, i.e., $x_2, x_3, \ldots, x_n$ can be found as functions of t.

It can be shown that the system of functions

$$x_1 = x_1(t), x_2 = x_2(t), \ldots, x_n = x_n(t)$$

constructed in this way is a solution to (19.9).

Example. Integrate the system

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = -x. \tag{19.13}$$

◄ Differentiating the first of these, we will get $\frac{d^2x}{dt^2} = \frac{dy}{dt}$, whence, using the second equation, we will have $\frac{d^2x}{dt^2} + x = 0$, i.e., a linear differential equation of the second order with constant coefficients and one unknown function. Its general solution has the form

$$x(t) = C_1 \cos t + C_2 \sin t.$$

From the first equation we find

$$y(t) = -C_1 \sin t + C_2 \cos t.$$

It is easy to verify that $x(t)$ and $y(t)$ at all C_1 and C_2 satisfy the given system.

Functions $x(t)$ and $y(t)$ can be represented in the form

$$x = A\sin(t + \alpha), \qquad y = A\cos(t + \alpha). \tag{19.14}$$

And so the integral curves of (19.13) are helical lines with lead $h = 2\pi$ and general axis $x = y = 0$, which is also an integral curve (Fig. 19.3).

Eliminating in (19.14) the parameter t, we obtain the equation $x^2 + y^2 = A^2$, since the phase paths of the system are circles with the centre at the origin of coordinates, i.e., projections of the helical lines on the xy-plane.

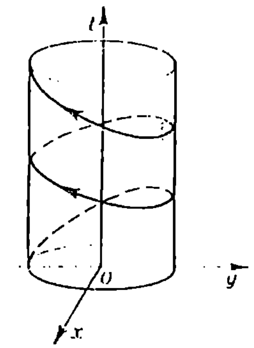


Fig. 19.3

At $A = 0$ the phase path consists of one point $x = 0$, $y = 0$, called the *stationary* (or *rest*) *point of the system.* ►

Remark. It may happen that the functions $x_2, x_3, \ldots, x_n$ cannot be found in terms of $t, x_1, \frac{dx_1}{dt}, \ldots, \frac{d^{n-1}x_1}{dt^{n-1}}$. We cannot, therefore, get the nth-order equation that is equivalent to the initial system. A simple example is the system

$$\frac{dx_1}{dt} = x_1, \qquad \frac{dx_2}{dt} = x_2.$$

An equivalent second-order equation in x_1 or x_2 cannot be found. This system is composed of a pair of first-order equations, each of which is independently integrable (yielding $x_1 = C_1e^t$ and $x_2 = C_2e^t$). ►

Normal systems

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \ldots, x_n), \qquad i = 1, 2, \ldots, n \tag{19.15}$$

are sometimes integrated by the *method of integrable combinations.*

An *integrable combination* is a differential equation that is a consequence of the system (19.15), but now readily integrable.

Let us integrate, for example, the system

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = x.$$

◄ Adding these equations together term by term yields the following integrable combinations:

$$\frac{d(x+y)}{dt} = x + y,$$

hence $x + y = C_1 e^t$.

Substracting term by term the second of these from the first one, we will obtain another integrable combination:

$$\frac{d(x-y)}{dt} = -(x-y),$$

hence $x - y = C_2 e^{-t}$. We have thus obtained two end results

$$x + y = C_1 e^t, \qquad x - y = C_2 e^{-t},$$

from which we can readily deduce the general solution of the system:

$$x(t) = \frac{1}{2}(C_1 e^t + C_2 e^{-t}), \qquad y(t) = \frac{1}{2}(C_1 e^t - C_2 e^{-t}). ►$$

One integrable combination yields one equation

$$\Phi_1(t, x_1, x_2, \ldots, x_n) = C_1,$$

that relates the independent variable t to the unknown functions x_1, $x_2, \ldots, x_n$. Such a finite equation is known as the *first integral* of the system (19.15). More precisely, the first integral of the system of differential equations (19.15) is a differentiable function $\Phi(t, x_1, x_2, \ldots, x_n)$ that is not identically constant but constant on any integral curve of the system.

If the first n integrals of system (19.15) have been found and all of them are independent, i.e., the Jacobian of the set of functions $\Phi_1, \ldots, \Phi_n$ is nonzero:

$$\begin{vmatrix} \dfrac{\partial\Phi_1}{\partial x_1} & \dfrac{\partial\Phi_1}{\partial x_2} & \cdots & \dfrac{\partial\Phi_1}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial\Phi_n}{\partial x_1} & \dfrac{\partial\Phi_n}{\partial x_2} & & \dfrac{\partial\Phi_n}{\partial x_n} \end{vmatrix} \neq 0,$$

then the problem of integration of system (19.15) is solved, since from the

system

$$\begin{cases} \Phi_1(t, x_1, x_2, \ldots, x_n) = C_1, \\ \Phi_2(t, x_1, x_2, \ldots, x_n) = C_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \Phi_n(t, x_1, x_2, \ldots, x_n) = C_n, \end{cases}$$

we find all the unknown functions $x_1(t), x_2(t), \ldots, x_n(t)$.

19.3 Systems of Linear Differential Equations

A system of differential equations is said to be *linear*, if it is linear in the unknown functions and their derivatives present in the equations of the system. A system of n linear first-order equations, written in normal form, is

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}(t)x_j + f_i(t) \qquad (i = 1, 2, \ldots, n); \tag{19.16}$$

in matrix form

$$\frac{d\mathbf{X}}{dt} = \mathbf{AX} + \mathbf{F}, \tag{19.17}$$

where

$$\mathbf{X} = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \qquad \mathbf{F} = \begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{pmatrix},$$

$$\mathbf{A} = \begin{pmatrix} a_{11}(t) & a_{12}(t) & & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & & a_{2n}(t) \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1}(t) & a_{n2}(t) & \ldots & a_{nn}(t) \end{pmatrix}.$$

Theorem 19.2. *If all functions $a_{ij}(t)$ and $f_i(t)$, $i, j = 1, 2, \ldots, n$, are continuous on the interval $a \leqslant t \leqslant b$, then in a sufficiently small neighbourhood of each point $M_0(t_0, x_1^0, x_2^0, \ldots, x_n^0)$, where $t_0 \in (a, b)$, the conditions of the theorem on the existence and uniqueness of the Cauchy problem are satisfied, and so through each such point passes only one integral curve of* (19.16).

In this event, the right-hand sides of (19.16) are continuous in the collection of arguments $t, x_1, x_2, \ldots, x_n$ and their partial derivatives with respect

to x_j, $j = 1, 2, \ldots, n$, are bounded, since these derivatives are equal to the coefficients $a_{ij}(t)$ that are continuous on the interval $[a, b]$.

We introduce the linear operator $L = d/dt - \mathbf{A}$.

System (19.17) then becomes

$$L[\mathbf{X}] = \mathbf{F}. \tag{19.18}$$

If F is a null matrix, i.e., $f_i(t) \equiv 0$, $i = 1, 2, \ldots, n$, in the interval (a, b), then system (19.17) is called a *linear homogeneous system* and has the form

$$L[\mathbf{X}] = 0. \tag{19.19}$$

We will give some theorems that establish the behaviour of solutions to linear systems.

Theorem 19.3. *If* $\mathbf{X}(t)$ *is a solution of a linear homogeneous system* $L[\mathbf{X}] = 0$, *then* $c\mathbf{X}(t)$, *where* c *is an arbitrary constant, is also a solution of the system.*

Theorem 19.4. *The sum* $\mathbf{X}_1(t) + \mathbf{X}_2(t)$ *of two solutions* $\mathbf{X}_1(t)$ *and* $\mathbf{X}_2(t)$ *of a homogeneous linear system of equations is a solution of the system.*

Corollary. A linear combination $\sum\limits_{i=1}^{m} c_i\mathbf{X}_i(t)$ with arbitrary constant coefficients c_i of the solutions $\mathbf{X}_1(t), \ldots, \mathbf{X}_m(t)$ of the linear homogeneous system of differential equations $L[\mathbf{X}] = 0$ is a solution of the system.

Theorem 19.5. *If* $\tilde{\mathbf{X}}(t)$ *is a solution of a linear inhomogeneous system* $L[\mathbf{X}] = \mathbf{F}$, *and* $\mathbf{X}_0(t)$ *is a solution of the corresponding homogeneous system* $L[\mathbf{X}] = 0$, *then the sum* $\tilde{\mathbf{X}}(t) + \mathbf{X}_0(t)$ *will be a solution of the inhomogeneous system* $L[\mathbf{X}] = \mathbf{F}$.

◄ As stated, $L[\tilde{\mathbf{X}}] \equiv \mathbf{F}$, $L[\mathbf{X}_0] \equiv 0$. Using the property of additivity of operator L, we obtain

$$L(\tilde{\mathbf{X}} + \mathbf{X}_0) = L[\tilde{\mathbf{X}}] + L[\mathbf{X}_0] \equiv \mathbf{F}.$$

This means that the sum $\tilde{\mathbf{X}}(t) + \mathbf{X}_0(t)$ is a solution of the inhomogeneous system $L[\mathbf{X}] = \mathbf{F}$.

Definition. Vectors $\mathbf{X}_1(t), \mathbf{X}_2(t), \ldots, \mathbf{X}_n(t)$, where

$$\mathbf{X}_k(t) = \begin{pmatrix} x_{1k}(t) \\ x_{2k}(t) \\ \vdots \\ x_{nk}(t) \end{pmatrix},$$

are called *linearly dependent* in the interval $a < t < b$, if there exist constant numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that

$$\alpha_1\mathbf{X}_1(t) + \alpha_2\mathbf{X}_2(t) + \ldots + \alpha_n\mathbf{X}_n(t) \equiv 0 \tag{19.20}$$

for $t \in (a, b)$, and at least one of α_i is nonzero. If (19.29) holds only at $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, then the vectors $\mathbf{X}_1(t), \mathbf{X}_2(t), \ldots, \mathbf{X}_n(t)$ are called *linearly independent* in (a, b).

It is worth noting that one vector identity (19.20) is equivalent to n identities

$$\sum_{k=1}^{n} \alpha_k x_{1k}(t) \equiv 0,$$
$$\sum_{k=1}^{n} \alpha_k x_{2k}(t) \equiv 0,$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$\sum_{k=1}^{n} \alpha_k x_{nk}(t) \equiv 0.$$

The determinant

$$W(t) = \begin{vmatrix} x_{11}(t) & x_{12}(t) & & x_{1n}(t) \\ x_{21}(t) & x_{22}(t) & & x_{2n}(t) \\ \vdots & \vdots & & \vdots \\ x_{n1}(t) & x_{n2}(t) & & x_{nn}(t) \end{vmatrix}$$

is called the *Wronskian* of the set of vectors $\mathbf{X}_1(t), \mathbf{X}_2(t), \ldots, \mathbf{X}_n(t)$.

Definition. Consider the linear homogeneous system

$$\frac{d\mathbf{X}}{dt} = \mathbf{A}(t)\mathbf{X}, \tag{19.21}$$

where $\mathbf{A}(t)$ is an $n \times n$ matrix with elements $a_{ij}(t)$. The system of n solutions $\mathbf{X}_1(t), \mathbf{X}_2(t), \ldots, \mathbf{X}_n(t)$ of the linear homogeneous system (19.21) which are linearly independent in the interval $a < t < b$ is called a *fundamental set*.

Theorem 19.6. *The Wronskian $W(t)$ of a fundamental set in the interval $a < t < b$ of solutions of system* (19.21), *whose coefficients $a_{ij}(t)$ are continuous on $a \leqslant t \leqslant b$, is nonzero at all points of the interval (a, b).*

Theorem 19.7 (on structure of general solution of a linear homogeneous system). *The general solution in the domain $a < t < b$, $|x_k| < +\infty$, $k = 1, 2, \ldots, n$, of the system* (19.21) *with coefficients $a_{ij}(t)$ continuous on the interval $a \leqslant t \leqslant b$ is a linear combination of n linearly independent in the interval $a < t < b$ solutions $\mathbf{X}_1(t), \mathbf{X}_2(t), \ldots, \mathbf{X}_n(t)$ of* (19.21):

$$\mathbf{X}_{0,0} = \sum_{i=1}^{n} c_i \mathbf{X}_i(t),$$

where $c_1, c_2, \ldots, c_n$ are arbitrary constants.

For example, the system

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -x_1$$

has, as is easily checked, the solutions

$$\mathbf{X}_1(t) = \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}, \qquad \mathbf{X}_2(t) = \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}.$$

These solutions are linearly independent, since the Wronskian is nonzero, i.e.,

$$W(t) = \begin{vmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{vmatrix} = 1.$$

The general solution of the system is

$$\mathbf{X}(t) = c_1\mathbf{X}_1(t) + c_2\mathbf{X}_2(t)$$

or

$$x_1(t) = c_1 \cos t + c_2 \sin t,$$
$$x_2(t) = -c_1 \sin t + c_2 \cos t,$$

where c_1, c_2 are arbitrary constants.

The square matrix

$$\mathscr{X}(t) = \begin{pmatrix} x_{11}(t) & x_{12}(t) & \dots & x_{1n}(t) \\ x_{21}(t) & x_{22}(t) & \dots & x_{2n}(t) \\ \vdots & \vdots & & \vdots \\ x_{n1}(t) & x_{n2}(t) & & x_{nn}(t) \end{pmatrix},$$

whose columns are linearly independent solutions $\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_n(t)$ of system (19.21), is known as the *fundamental matrix* of the system. It can easily be verified that the fundamental matrix obeys the matrix equation

$$\frac{d\mathscr{X}}{dt} = \mathbf{A}(t)\mathscr{X}(t).$$

If $\mathscr{X}(t)$ is the fundamental matrix of (19.21), then the general solution of the system can be represented as

$$\mathbf{X}(t) = \mathscr{X}(t)\mathbf{C}, \tag{19.22}$$

where $\mathbf{C} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$ is a constant column matrix with arbitrary elements.

Setting in (19.22) $t = t_0$ produces

$$\mathbf{X}(t_0) = \mathcal{X}(t_0)\mathbf{C}, \qquad \text{hence} \qquad \mathbf{C} = \mathcal{X}^{-1}(t_0)\,\mathbf{X}(t_0).$$

Therefore,

$$\mathbf{X}(t) = \mathcal{X}(t)\mathcal{X}^{-1}(t_0)\mathbf{X}(t_0).$$

The matrix $\mathcal{X}(t)\mathcal{X}^{-1}(t_0) = \mathbf{K}(t, t_0)$ is called the *Cauchy matrix*. In terms of the Cauchy matrix the solution of (19.21) can be represented as

$$\mathbf{X}(t) = \mathbf{K}(t, t_0)\mathbf{X}(t_0). \tag{19.23}$$

Theorem 19.8 (on structure of general solution of a linear inhomogeneous system). *The general solution in the domain $a < t < b$, $|x_k| < +\infty$, $k = 1, 2, \ldots, n$, of the linear system*

$$\frac{d\mathbf{X}}{dt} = \mathbf{A}(t)\,\mathbf{X} + \mathbf{F}(t) \tag{19.24}$$

with coefficients $a_{ij}(t)$ and right-hand sides $f_i(t)$ continuous for $a \leqslant t \leqslant b$, is the sum of the general solution $\sum_{k=1}^{n} c_k\mathbf{X}_k(t)$ of the corresponding homogeneous system and some particular solution $\tilde{\mathbf{X}}(t)$ of the inhomogeneous system (19.24):

$$\mathbf{X}_{\text{g.i.}} = \mathbf{X}_{\text{g.h.}} + \mathbf{X}_{\text{p.i.}}.$$

Given the general solution of the linear homogeneous system (19.21), a particular solution of the inhomogeneous system can be found by variation of constants (Lagrange method). Indeed, let

$$\mathbf{X}(t) = \sum_{k=1}^{n} c_k\mathbf{X}_k(t)$$

be the general solution of (19.21), then

$$\frac{d\mathbf{X}_k}{dt} \equiv \mathbf{A}(t)\,\mathbf{X}_k(t), \qquad t \in (a, b) \qquad (k = 1, 2, \ldots, n).$$

Here $\mathbf{X}_k(t)$ are linearly independent.

We will seek the particular solution of the inhomogeneous system (19.24) in the form

$$\tilde{\mathbf{X}}(t) = \sum_{k=1}^{n} c_k(t)\,\mathbf{X}_k(t), \tag{19.25}$$

where $c_k(t)$, $k = 1, 2, \ldots, n$, are unknown functions of t. Differentiating

$\tilde{X}(t)$ with respect to t gives

$$\frac{d\tilde{X}}{dt} = \sum_{k=1}^{n} c_k(t)\frac{dX_k}{dt} + \sum_{k=1}^{n} c_k'(t)X_k(t).$$

Substituting $\tilde{X}(t)$ and $d\tilde{X}/dt$ into (19.24) yields

$$\sum_{k=1}^{n} c_k(t)\left(\frac{dX_k}{dt} - A(t)X_k(t)\right) + \sum_{k=1}^{n} c_k'(t)X_k(t) = F(t).$$

Since $dX_k/dt - A(t)X_k(t) \equiv 0$, then we will determine $c_k'(t)$ from

$$\sum_{k=1}^{n} c_k'(t)X_k(t) = F(t),$$

or, if we expand it,

$$\begin{cases} c_1'(t)x_{11}(t) + c_2'(t)x_{12}(t) + \ldots + c_n'(t)x_{1n}(t) = f_1(t) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_1'(t)x_{n1}(t) + c_2'(t)x_{n2}(t) + \ldots + c_n'(t)x_{nn}(t) = f_n(t). \end{cases} \tag{19.26}$$

System (19.26) is a linear algebraic system in $c_k'(t)$, whose determinant is the Wronskian $W(t)$ of the fundamental set of solutions $X_1(t), X_2(t), \ldots, X_n(t)$. This determinant is nonzero everywhere in the interval $a < t < b$, since (19.26) has the only solution

$$c_k'(t) = \varphi_k(t) \qquad (k = 1, 2, \ldots, n),$$

where $\varphi_k(t)$ are known continuous functions. Integrating these gives

$$c_k(t) = \int \varphi_k(t)\,dt \qquad (k = 1, 2, \ldots, n).$$

Substituting these values of $c_k(t)$ into (19.25), we find the particular solution of (19.24)

$$\tilde{X}(t) = \sum_{k=1}^{n} X_k(t)\int \varphi_k(t)dt.$$

Here $\int \varphi_k(t)dt$ is assumed to be one of the antiderivatives for $\varphi_k(t)$.

19.4 Systems of Linear Differential Equations with Constant Coefficients

Consider the linear system of differential equations

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}x_j + f_i(t) \qquad (i = 1, 2, \ldots, n),$$

where all a_{ij} $(i, j = 1, 2, \ldots, n)$ are constants. Such a system can be integrated by reduction to one equation of a higher order, which will also be linear and with constant coefficients. Another effective method of integrating systems with constant coefficients is the Laplace method.

We will consider the *Euler method* of integrating linear homogeneous systems of differential equations with constant coefficients. We will seek the solution of the system

$$\begin{cases} \dfrac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ \dfrac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n \end{cases} \tag{19.27}$$

in the form

$$x_1 = \alpha_1 e^{\lambda t}, \quad x_2 = \alpha_2 e^{\lambda t}, \quad \ldots, \quad x_n = \alpha_n e^{\lambda t}, \tag{19.28}$$

where λ, α_1, α_2, $\ldots$, α_n are constants. Substituting x_k in the form (19.28) into (19.27), cancelling out $e^{\lambda t}$, and gathering all the terms on one side of the equality, we will arrive at

$$\begin{cases} (a_{11} - \lambda)\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n = 0, \\ a_{21}\alpha_1 + (a_{22} - \lambda)\alpha_2 + \ldots + a_{2n}\alpha_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\alpha_1 + a_{n2}\alpha_2 + \ldots + (a_{nn} - \lambda)\alpha_n = 0. \end{cases} \tag{19.29}$$

For system (19.29) of linear homogeneous algebraic equations with n unknowns $\alpha_1, \alpha_2, \ldots, \alpha_n$ to have a nontrivial solution, it is necessary and sufficient that its determinant

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{19.30}$$

Equation (19.30) is called the *characteristic equation*. Its left-hand side contains a polynomial in λ to power n. From this equation we determine those values of λ at which (19.22) has nontrivial solutions $\alpha_1, \alpha_2, \ldots, \alpha_n$. If all the roots λ_i, $i = 1, 2, \ldots, n$, of the characteristic equation (19.30) are different, then substituting them one by one into (19.29) gives the corresponding nontrivial solutions $\alpha_{1i}, \alpha_{2i}, \ldots, \alpha_{ni}$, $i = 1, 2, \ldots, n$, of the

system. We thus find n solutions of the original system of differential equations (19.27) in the form

$$x_{1i} = \alpha_{1i}e^{\lambda_i t},\ x_{2i} = \alpha_{2i}e^{\lambda_i t},\ \ldots,\ x_{ni} = \alpha_{ni}e^{\lambda_i t}\ (i = 1, 2, \ldots, n), \quad (19.31)$$

where the second subscript is the number of the solution, and the first one is the number of the unknown function. For the linear homogeneous system (19.27) we thus form n particular solutions

$$\mathbf{X}_1 = \begin{pmatrix} x_{11}(t) \\ x_{21}(t) \\ \vdots \\ x_{n1}(t) \end{pmatrix}, \qquad \mathbf{X}_2(t) = \begin{pmatrix} x_{12}(t) \\ x_{22}(t \\ \vdots \\ x_{n2}(t) \end{pmatrix}, \qquad \ldots,$$

$$\mathbf{X}_n(t) = \begin{pmatrix} x_{1n}(t) \\ x_{2n}(t) \\ \vdots \\ x_{nn}(t) \end{pmatrix}, \qquad (19.82)$$

which, as can readily be verified, form the fundamental set of solutions of (19.27).

In consequence, the general solution of (19.27) has the form

$$\mathbf{X}(t) = c_1\mathbf{X}_1(t) + c_2\mathbf{X}_2(t) + \ldots + c_n\mathbf{X}_n(t),$$

or

$$\begin{aligned} x_1(t) &= c_1x_{11}(t) + c_2x_{12}(t) + \ldots + c_nx_{1n}(t), \\ x_2(t) &= c_1x_{21}(t) + c_2x_{22}(t) + \ldots + c_nx_{2n}(t), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n(t) &= c_1x_{n1}(t) + c_2x_{n2}(t) + \ldots + c_nx_{nn}(t), \end{aligned}$$

where $c_1, c_2, \ldots, c_n$ are arbitrary constants.

We will not consider here the case where a characteristic equation has multiple roots.

Example. Solve the system

$$\frac{dx_1}{dt} = -x_1 + 2x_2, \qquad \frac{dx_2}{dt} = 2x_1 - x_2.$$

◄ We will look for the solution in the form $x_1 = \alpha_1e^{\lambda t}$, $x_2 = \alpha_2e^{\lambda t}$. The characteristic equation

$$\begin{vmatrix} -1-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix} = 0 \quad \text{or} \quad \lambda^2 + 2\lambda - 3 = 0$$

has the roots $\lambda_1 = 1$ and $\lambda_2 = -3$.

The system (19.29) for α_1 and α_2 looks like

$$\begin{cases}(-1-\lambda)\alpha_1 + 2\alpha_2 = 0,\\ 2\alpha_1 + (-1-\lambda)\alpha_2 = 0.\end{cases} \tag{*}$$

Substituting $\lambda = 1$ into (*) gives

$$\begin{cases}-2\alpha_{11} + 2\alpha_{21} = 0,\\ \ \ 2\alpha_{11} - 2\alpha_{21} = 0,\end{cases}$$

hence $\alpha_{21} = \alpha_{11}$. Thus

$$x_{11} = \alpha_{11}e^t, \qquad x_{21} = \alpha_{11}e^t.$$

Putting in (*) $\lambda = -3$, we find $\alpha_{22} = -\alpha_{12}$, therefore

$$x_{12} = \alpha_{12}e^{-3t}, \qquad x_{22} = -\alpha_{12}e^{-3t}.$$

The general solution of the system is

$$x_1(t) = c_1\alpha_{11}e^t + c_2\alpha_{12}e^{-3t},$$
$$x_2(t) = c_1\alpha_{11}e^t - c_2\alpha_{12}e^{-3t},$$

or

$$x_1(t) = C_1e^t + C_2e^{-3t},$$
$$x_2(t) = C_1e^t - C_2e^{-3t}. \blacktriangleright$$

We now turn to the matrix method of integration of (19.16). We rewrite it in the form

$$\frac{d\mathbf{X}}{dt} = \mathbf{AX}, \tag{19.33}$$

where

$$\mathbf{X}(t) = \begin{pmatrix} x_1(t)\\ x_2(t)\\ \vdots\\ x_n(t)\end{pmatrix}$$

and $\mathbf{A} = (a_{ij})_{i,j=1}^n$ is an $n \times n$ matrix with constant real elements a_{ij}.

We recall some concepts of linear algebra. Vector $\mathbf{g} \neq 0$ is called an eigenvector of matrix $\mathbf{A}$, if

$$\mathbf{Ag} = \lambda\mathbf{g}.$$

Number λ is called the eigenvalue of $\mathbf{A}$ that corresponds to the eigenvector $\mathbf{g}$ and is a root of the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0,$$

where $\mathbf{I}$ is the identity matrix.

We assume that all the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $\mathbf{A}$ are different. In that case, the eigenvectors $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n$ are linearly independent and there exists an $n \times n$ matrix $\mathbf{T}$ that reduces $\mathbf{A}$ to diagonal form, i.e., such that

$$\Lambda = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}. \tag{19.34}$$

The columns of $\mathbf{T}$ are the coordinates of the eigenvectors $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n$ of $\mathbf{A}$.

We introduce the following concepts. Let $\mathbf{B}(t)$ be an $n \times n$ matrix whose elements $b_{ij}(t)$ are functions of t defined on the set Ω. Matrix $\mathbf{B}(t)$ is said to be continuous on Ω, if all of its elements $b_{ij}(t)$ are continuous on Ω. Matrix $\mathbf{B}(t)$ is said to be differentiable on Ω, if all the elements $b_{ij}(t)$ of the matrix are differentiable on Ω. The derivative $d\mathbf{B}(t)/dt$ of matrix $\mathbf{B}(t)$ is a matrix whose elements are the derivatives $db_{ij}(t)/dt$ of the corresponding elements of $\mathbf{B}(t)$.

Let $\mathbf{B}(t)$ be an $n \times n$ matrix, and $\mathbf{X}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}$ be a column vector.

Using the rules of matrix algebra, we can make sure by direct check that the following formula holds:

$$\frac{d}{dt}(\mathbf{B}(t)\mathbf{X}(t)) = \frac{d\mathbf{B}(t)}{dt}\mathbf{X}(t) + \mathbf{B}(t)\frac{d\mathbf{X}}{dt}. \tag{19.35}$$

Specifically, if $\mathbf{B}$ is a constant matrix, then $\frac{d}{dt}(\mathbf{B}\mathbf{X}(t)) = \mathbf{B}\frac{d\mathbf{X}}{dt}$, since $\frac{d\mathbf{B}}{dt}$ is the null matrix.

Theorem 19.9. *If the eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *of a matrix* $\mathbf{A}$ *are different, then the general solution of the system* (19.33) *has the form*

$$\mathbf{X}(t) = c_1 e^{\lambda_1 t}\mathbf{g}_1 + c_2 e^{\lambda_2 t}\mathbf{g}_2 + \ldots + c_n e^{\lambda_n t}\mathbf{g}_n, \tag{19.36}$$

where $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n$ are the column eigenvectors of $\mathbf{A}$, and $c_1, c_2, \ldots, c_n$ are arbitrary constants.

◄ We introduce a new unknown column vector $\mathbf{Y}(t)$ by

$$\mathbf{X}(t) = \mathbf{T}\mathbf{Y}(t), \tag{19.37}$$

where $\mathbf{T}$ is a matrix that reduces $\mathbf{A}$ to diagonal form. Substituting $\mathbf{X}(t)$

from (19.37) into (19.33) yields

$$\mathbf{T}\frac{d\mathbf{Y}}{dt} = \mathbf{ATY}.$$

Premultiplying both sides of this by $\mathbf{T}^{-1}$ and considering that $\mathbf{T}^{-1}\mathbf{AT} = \Lambda$, we arrive at

$$\frac{d\mathbf{Y}}{dt} = \Lambda\mathbf{Y},$$

or

$$\begin{cases} \dfrac{dy_1}{dt} = \lambda_1 y_1, \\ \dfrac{dy_2}{dt} = \lambda_2 y_2, \\ \dots\dots\dots \\ \dfrac{dy_n}{dt} = \lambda_n y_n. \end{cases} \tag{19.38}$$

We have thus obtained a system of n independent equations which is easily integrable

$$y_1 = c_1 e^{\lambda_1 t}, \qquad y_2 = c_2 e^{\lambda_2 t}, \ \dots, \ y_n = c_n e^{\lambda_n t}.$$

Here $c_1, c_2, \dots, c_n$ are arbitrary constants.

Introducing unit n-dimensional column vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

we can represent the solution $\mathbf{Y}(t)$ as

$$\mathbf{Y}(t) = c_1 e^{\lambda_1 t}\mathbf{e}_1 + c_2 e^{\lambda_2 t}\mathbf{e}_2 + \dots + c_n e^{\lambda_n t}\mathbf{e}_n. \tag{19.39}$$

By (19.37), we have $\mathbf{X}(t) = \mathbf{TY}(t)$. Since the columns of $\mathbf{T}$ are eigenvectors of $\mathbf{A}$, then $\mathbf{Te}_k = \mathbf{g}_k$, where $\mathbf{g}_k$ is the kth eigenvector of $\mathbf{A}$. Therefore, substituting (19.39) into (19.37), we will obtain (19.36). ►

Thus, if the matrix $\mathbf{A}$ of the system (19.33) has different eigenvalues, to obtain the general solution of the system it is necessary:

(1) to find the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of the matrix as roots of the algebraic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$;

(2) to find all the eigenvectors $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$;

(3) to write the general solution of (19.33) using (19.36).

Example. Solve the system

$$\begin{cases} \dfrac{dx}{dt} = 3x + y, \\ \dfrac{dy}{dt} = 2x + 2y. \end{cases}$$

◄ Matrix $\mathbf{A}$ of the system has the form

$$\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}.$$

(1) We set up the characteristic equation

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 3-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} = 0$$

or

$$\lambda^2 - 5\lambda + 4 = 0.$$

The roots of the equation are $\lambda_1 = 4$ and $\lambda_2 = 1$.

(2) We now find the eigenvectors

$$\mathbf{g}_1 = \begin{pmatrix} g_{11} \\ g_{12} \end{pmatrix}, \quad \mathbf{g}_2 = \begin{pmatrix} g_{21} \\ g_{22} \end{pmatrix}.$$

For $\lambda = 4$ the system becomes

$$\begin{cases} -g_{11} + g_{12} = 0, \\ 2g_{11} - 2g_{12} = 0, \end{cases}$$

hence $g_{11} = g_{12}$, since $g_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

For $\lambda = 1$ we similarly find $\mathbf{g}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

(3) Using (19.36), we obtain the general solution of the system

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^{4t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{t} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

or

$$x(t) = c_1 e^{4t} + c_2 e^{t},$$
$$y(t) = c_1 e^{4t} - 2c_2 e^{t}. \text{ ►}$$

A characteristic equation may have real and complex roots. Since the coefficients a_{ij} of the system (19.33) are assumed to be real, then the characteristic equation $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$ will have real coefficients. Therefore, along with the complex root λ it will have the root λ^*, which is a complex conjugate to λ. It can easily be shown that if $\mathbf{g}$ is the eigenvector corresponding to the eigenvalue λ, then λ^* will also be the eigenvalue to which corresponds the eigenvector $\mathbf{g}^*$, which is a complex conjugate to $\mathbf{g}$.

When λ is complex, the solution $X = e^{\lambda t}g$ of (19.33) will also be complex. The real part $X_1 = \mathrm{Re}\,(e^{\lambda t}g)$ and the imaginary part $X_2 = \mathrm{Im}\,(e^{\lambda t}g)$ of this solution will again be solutions of system (19.33). Corresponding to the eigenvalue λ^* will be a pair of real solutions X_1 and $-X_2$, i.e., the same pair as for λ. And so corresponding to the pair λ and λ^* of complex conjugate eigenvalues is a pair of real solutions of (19.33).

Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be real eigenvalues and $\lambda_{k+1}, \lambda_{k+1}^*, \lambda_{k+2}, \lambda_{k+2}^*, \ldots$ be complex eigenvalues.

Any real solution of (19.33) will then have the form

$$X(t) = c_1 e^{\lambda_1 t}g_1 + c_2 e^{\lambda_2 t}g_2 + \ldots + c_k e^{\lambda_k t}g_k$$
$$+ c_{k+1}\mathrm{Re}\,(e^{\lambda_{k+1}t}g_{k+1}) + c_{k+2}\mathrm{Im}\,(e^{\lambda_{k+1}t}g_{k+1}) + \ldots,$$

where c_i are arbitrary constants.

Example. Solve the system

$$\begin{cases} \dfrac{dx}{dt} = x - 3y, \\ \dfrac{dy}{dt} = 3x + y. \end{cases}$$

◄ The corresponding matrix is

$$A = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}.$$

(1) The characteristic equation is

$$\begin{vmatrix} 1-\lambda & -3 \\ 3 & 1-\lambda \end{vmatrix} = 0$$

or

$$(1-\lambda)^2 + 9 = 0.$$

Its roots are $\lambda_1 = 1 + 3i$ and $\lambda_2 = 1 - 3i$.

(2) The eigenvectors of the matrix are

$$g_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad g_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

(3) The solution of the system is

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = a_1 e^{(1+3i)t}\begin{pmatrix} 1 \\ -i \end{pmatrix} + a_2 e^{(1-3i)t}\begin{pmatrix} 1 \\ i \end{pmatrix},$$

where a_1 and a_2 are arbitrary complex constants.

We now wish to find the real solutions of the system. Using the Euler formula $e^{i\varphi} = \cos\varphi + i\sin\varphi$, we get

$$e^{3it}\begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} \cos 3t \\ \sin 3t \end{pmatrix} + i\begin{pmatrix} \sin 3t \\ -\cos 3t \end{pmatrix}.$$

Any real solution will thus be

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = c_1 e^t \begin{pmatrix} \cos 3t \\ \sin 3t \end{pmatrix} + c_2 e^t \begin{pmatrix} \sin 3t \\ -\cos 3t \end{pmatrix}$$

or

$$x(t) = c_1 e^t \cos 3t + c_2 e^t \sin 3t,$$
$$y(t) = c_1 e^t \sin 3t - c_2 e^t \cos 3t,$$

where c_1 and c_2 are arbitrary real numbers. ►

Exercises

Integrate the following systems by elimination:

1. $\begin{cases} \dfrac{dx}{dt} = -x + 2y, \\ \dfrac{dy}{dt} = 2x - y. \end{cases}$ 2. $\begin{cases} \dfrac{dx}{dt} = 3x - 2y, \\ \dfrac{dy}{dt} = 2x - y. \end{cases}$ 3. $\begin{cases} \dfrac{d^2x}{dt^2} = y, \\ \dfrac{d^2y}{dt^2} = x. \end{cases}$

Integrate the following systems by the method of integrable combinations:

4. $\begin{cases} \dfrac{dx}{dt} = x + y, \\ \dfrac{dy}{dt} = x + y + t. \end{cases}$ 5. $\begin{cases} \dfrac{dx}{dt} = \dfrac{1}{y}, \\ \dfrac{dy}{dt} = \dfrac{1}{x}. \end{cases}$

Integrate the following systems by the matrix method:

6. $\begin{cases} \dfrac{dx}{dt} = 2x + y, \\ \dfrac{dy}{dt} = 3x + 4y. \end{cases}$ 7. $\begin{cases} \dfrac{dx}{dt} = x - y, \\ \dfrac{dy}{dt} = x + y. \end{cases}$

Answers

1. $x = C_1 e^t + C_2 e^{-3t}$, $y = C_1 e^t - C_2 e^{-3t}$. 2. $x = \left(C_1 + \dfrac{C_2}{2}\right) e^t + C_2 t e^t$, $y = C_1 e^t + C_2 t e^t$. 3. $x = C_1 e^t + C_2 e^{-t} + C_3 \sin t + C_4 \cos t$, $y = C_1 e^t + C_2 e^{-t} - C_3 \sin t - C_4 \cos t$. 4. $x = C_1 + C_2 e^{2t} - \dfrac{t^2}{4} - \dfrac{t}{4} - \dfrac{1}{8}$, $y = -C_1 + C_2 e^{2t} + \dfrac{t^2}{4} - \dfrac{t}{4} - \dfrac{1}{8}$. 5. $x = \sqrt{C_2(2t + C_1)}$, $y = \sqrt{\dfrac{2t + C_1}{C_2}}$. 6. $x = C_1 e^t + C_2 e^{5t}$, $y = -C_1 e^t + 3C_2 e^{5t}$. 7. $x = C_1 e^t \cos t + C_2 e^t \sin t$, $y = C_1 e^t \sin t - C_2 e^t \cos t$.

Chapter 20

Stability Theory

20.1 Preliminaries

Let us now look at how solutions to the Cauchy problem depend on the initial data.

Consider the Cauchy problem

$$\frac{dx}{dt} = f(t, x), \tag{20.1}$$

$$x(t_0) = x_0. \tag{20.2}$$

If the function $f(t, x)$ is continuous in the collection of arguments and has a bounded derivative $\partial f/\partial x$ in a domain Ω of t, x that contains a point (t_0, x_0), then there exists a unique solution to the Cauchy problem (20.1)-(20.2). Changing the values of t_0 and x_0 will also change the solution. One question important for applications is: just how will this affect the solution? This is also a question of principle. If some physics problem comes down to the Cauchy problem then the initial values are to be found from experiment and we cannot vouch for the precision of the measurement. And if arbitrarily small changes in initial data are able to drastically change the solution, then the mathematical model will hardly be suitable to describe the real process.

The continuous dependence of the solution on the initial conditions is established by

Theorem 20.1. *If the right-hand side $f(t, x)$ of the differential equation* (20.1) *is continuous in the collection of variables and has a bounded partial derivative $\partial f/\partial x$ in a domain G of t, x, then the solution $x(t) = x(t;\, t_0, x_0)$ that satisfies the initial condition $x(t_0) = x_0$, where $(t_0, x_0) \in G$, is continuously dependent on the initial data.*

In other words, suppose that through a point (t_0, x_0) passes the solution $x(t)$ of (20.1) defined on the interval $\alpha \leqslant t \leqslant \beta$, $t_0 \in (\alpha, \beta)$. Then, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for $|\tilde{t}_0 - t_0| < \delta$, $|\tilde{x}_0 - x_0| < \delta$ the solution $\tilde{x}(t)$ of (20.1) that passes through $(\tilde{t}_0, \tilde{x}_0)$ exists on $[\alpha, \beta]$ and differs there from $x(t)$ by less than ε:

$$|x(t) - \tilde{x}(t)| < \varepsilon \qquad \forall t \in [\alpha, \beta].$$

A similar theorem is valid also for the system of differential equations

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \dots, x_n) \qquad (i = 1, 2, \dots, n).$$

If the conditions of Theorem 20.1 are met, there exists a solution to the Cauchy problem that is unique and continuously dependent on the initial conditions. The Cauchy problem is then said to be *correctly* formulated. It is of significance that the interval $[\alpha, \beta]$ of t is finite. In many problems, however, we are interested in the dependence of the solution on the initial data in the infinite interval $t_0 \leqslant t < +\infty$. The transition from a finite interval in which we consider the continuous dependence of the solution on the initial conditions to an infinite one changes dramatically the nature of the problem and examination procedures. The problem comes under the heading of stability theory due to A. Lyapunov.

We will now sketch the idea of the *extendibility of solutions*. Suppose we have a system

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \dots, x_n) \qquad (i = 1, 2, \dots, n), \tag{20.3}$$

where t is an independent variable (time), $x_1(t), x_2(t), \dots, x_n(t)$ are the desired functions; $f_i(t, x_1, x_2, \dots, x_n)$ are functions defined for $t \in (a, +\infty)$ and $x_1, x_2, \dots, x_n$ from a domain $D \subset R^n$. If $f_i(t, x_1, x_2, \dots, x_n)$, $i = 1, 2, \dots, n$, are continuous in their domain in the collection of arguments and have bounded partial derivatives in $x_1, x_2, \dots, x_n$, then for (20.3) holds the *local* theorem of existence: for each set of values $(t_0, x_1^0, x_2^0, \dots, x_n^0)$, $t_0 \in (a, +\infty)$, $(x_1^0, x_2^0, \dots, x_n^0) \in D$ there exists the unique solution $x_1(t)$, $x_2(t), \dots, x_n(t)$ of (20.3) defined in an interval $(t_0 - h_0, t_0 + h_0) \subset (a, +\infty)$ of t and subject to the initial conditions

$$x_i(t_0) = x_i^0 \qquad (i = 1, 2, \dots, n). \tag{20.4}$$

We introduce the following concepts. Let $x_1(t), x_2(t), \dots, x_n(t)$ be a solution of the Cauchy problem (20.3)-(20.4) defined in an interval $I = (t_1, t_2)$. This solution can be extended to a larger time interval.

The solution $y_1(t), y_2(t), \dots, y_n(t)$ is called the *extension of the solution* $x_1(t), x_2(t), \dots, x_n(t)$, if it is defined in a larger interval $I_1 \supset I$ and coincides with $x_1(t), x_2(t), \dots, x_n(t)$ when $t \in I$.

A solution is called *infinitely extendible* (*infinitely extendible to the right or left*), if it can be extended to the entire axis $-\infty < t < +\infty$ (to the half-axes $t_0 \leqslant t < +\infty$ or $-\infty < t \leqslant t_0$, respectively).

Later in the chapter we will need to know whether there exists the solution $x_i(t)$, $i = 1, 2, \dots, n$, for $t_0 \leqslant t < +\infty$ (global existence theorem).

The property is inherent in the linear system

$$\frac{dx_i}{dt} = \sum_{i=1}^{n} a_{ij}(t)x_j + f_i(t) \qquad (i = 1, 2, \ldots, n),$$

where $a_{ij}(t)$ and $f_i(t)$ are continuous functions on $[t_0, +\infty)$. Each of its solutions $x_i(t)$, $i = 1, 2, \ldots, n$, exists on $[t_0, +\infty)$ (infinitely extendible to the right) and is unique. Not all systems have such a property.

For example, for the scalar equation

$$\frac{dx}{dt} = x^2 \tag{20.5}$$

the function $f(t, x) \equiv x^2$ is continuous and has derivatives of all orders in x. It can easily be checked that the function $x = \alpha/(1 - \alpha t)$ is a solution of the problem

$$\frac{dx}{dt} = x^2, \qquad x(0) = \alpha, \qquad \alpha > 0.$$

However, this solution only exists in the interval $(-\infty, 1/\alpha)$, which depends on the initial conditions, and is not extendible to the interval $(-\infty, 1/\alpha]$.

Equation (20.5) is the equation of superfast multiplication, when the growth of a population is proportional to the number of all pairs possible. Its solution shows that within a short span of time the population increases beyond limits (whereas the actual growth law is exponential).

Problem. Show that the solution to the equation $dx/dt = x^2 + 1$ cannot be extended in an infinite manner to the right or left.

20.2 Stability in the Sense of Lyapunov. Basic Concepts and Definitions

Consider the differential equation of the first order

$$\frac{dx}{dt} = f(t, x), \tag{20.1}$$

where $f(t, x)$ is defined and continuous for $t \in (a, +\infty)$ and x from a domain D and has a bounded partial derivative $\partial f/\partial x$. Let a function $x = \varphi(t)$ be a solution of equation (20.1) that meets the initial condition $x|_{t=t_0} = \varphi(t_0)$, $t_0 > a$. Suppose that $x = x(t)$ is a solution of the same equation but satisfying another initial condition $x|_{t=t_0} = x(t_0)$. It is also supposed that the solutions $\varphi(t)$ and $x(t)$ are defined for all $t \geqslant t_0$, i.e., they are extendible without limit to the right.

Definition. A solution $x = \varphi(t)$ of (20.1) is said to be *stable in the sense of Lyapunov*, as $t \to +\infty$, if for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for any solution $x = x(t)$ of the equation from the inequality

$$|x(t_0) - \varphi(t_0)| < \delta \tag{20.6}$$

follows the inequality

$$|x(t) - \varphi(t)| < \varepsilon \tag{20.7}$$

for all $t \geqslant t_0$ (we can always assume that $\delta \leqslant \varepsilon$).

This means that all solutions whose initial values are close to those of the solution $x = \varphi(t)$ remain close also at all $t \geqslant t_0$.

Geometrically, this means the following. The solution $x = \varphi(t)$ of (20.1) is stable if, however narrow an ε-band containing the curve $x = \varphi(t)$, all of the integral curves $x = x(t)$ of the equation that are sufficiently close to it at the initial moment of time $t = t_0$ are all contained in the ε-band at all $t \geqslant t_0$ (Fig. 20.1).

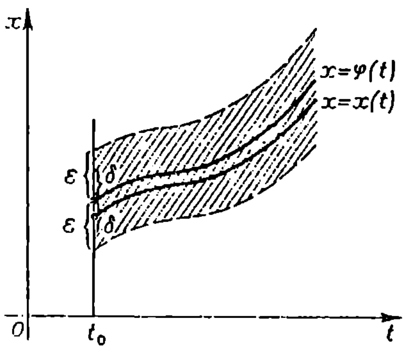


Fig. 20.1

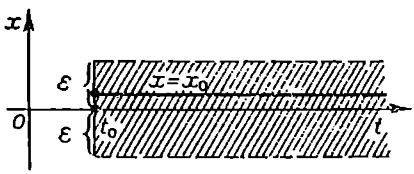


Fig. 20.2

If for arbitrarily small $\delta > 0$ at least for one solution $x = x(t)$ of (20.1) the inequality (20.7) does not hold, then the solution $x = \varphi(t)$ of the equation is said to be *unstable*. We should also regard as unstable a solution inextendible to the right as $t \to +\infty$.

Definition. A solution $x = \varphi(t)$ of (20.1) is said to be *asymptotically stable* if

(1) it is stable;

(2) there exists $\delta_1 > 0$ such that for any solution $x = x(t)$ of equation (20.1) satisfying the condition $|x(t_0) - \varphi(t_0)| < \delta_1$, we have

$$\lim_{t \to +\infty} |x(t) - \varphi(t)| = 0.$$

This means that all solutions whose initial conditions are close to the asympotincludes stable solution $x = \varphi(t)$ not only remain close to it for $t \geqslant t_0$, but also approach it without bound as $t \to +\infty$. Consider a simple physical model: a ball lies at the bottom of a hemispherical depression (in the equilibrium position). If we slightly disturb the ball from its equilibrium, it will swing about it. Without friction the equilibrium will be stable, with friction the oscillations of the ball will decrease with time, i.e., the equilibrium will be asymptotically stable.

Examples. (1) Examine for stability the trivial solution $x \equiv 0$ of the equation

$$\frac{dx}{dt} = 0. \tag{*}$$

◄ The solution $x \equiv 0$ will clearly meet the initial condition $x|_{t=t_0} = 0$. A solution of (*) that satisfies the initial condition $x|_{t=t_0} = x_0$ has the form $x \equiv x_0$. It is easily seen (Fig. 20.2) that, whatever the ε-band around the integral curve $x = 0$, there exists $\delta > 0$, e.g., $\delta = \varepsilon$, such that any integral curve $x = x_0$ for which $|x_0 - 0| < \delta$ wholly lies within the ε-band for all $t \geqslant t_0$. Hence the solution $x \equiv 0$ is stable. There is no asymptotic stability since the straight line $x = x_0$ does not tend to the line $x = 0$ as $t \to +\infty$. ►

(2) Examine for stability the trivial solution $x \equiv 0$ of the equation

$$\frac{dx}{dt} = -a^2x \qquad (a = \text{const}). \tag{**}$$

◄ The solution of equation (**) with the initial condition $x_{t=t_0} = x_0$ has the form

$$x = x_0e^{-a^2(t-t_0)}.$$

Take any $\varepsilon > 0$ and consider the difference of the solutions $x(t)$ and $\varphi(t) \equiv 0$:

$$x(t) - \varphi(t) = x_0e^{-a^2(t-t_0)} - 0 = (x_0 - 0)e^{-a^2(t-t_0)}. \tag{***}$$

Since $e^{-a^2(t-t_0)} \leqslant 1$ for all $t \geqslant t_0$, it follows from (***) that there exists $\delta > 0$, e.g., $\delta = \varepsilon$, such that for $|x_0 - 0| < \delta = \varepsilon$ we have

$$|x(t) - \varphi(t)| = |x_0 - 0|e^{-a^2(t-t_0)} < \varepsilon \ \forall t \geqslant t_0.$$

By definition this implies that the solution $\varphi(t) \equiv 0$ of (**) is stable.

Besides, we have

$$\lim_{t\to+\infty} |x(t) - \varphi(t)| = \lim_{t\to+\infty} |x_0|e^{-a^2(t-t_0)} = 0,$$

therefore, the solution $\varphi(t) \equiv 0$ is asymptotically stable (Fig. 20.3). ►

(3) Show that the solution $\varphi(t) \equiv 0$ of the equation

$$\frac{dx}{dt} = a^2 x$$

is unstable.

◄ For arbitrarily small $|x_0|$ the solution $x(t) = x_0 e^{a^2(t-t_0)}$ of the equation does not satisfy the condition $|x(t) - 0| = |x_0| e^{a^2(t-t_0)} < \varepsilon$ for sufficiently large $t > t_0$. Further, for any $x_0 \neq 0$ we have $|x(t)|_{t \to +\infty} \to +\infty$ (Fig. 20.4). ►

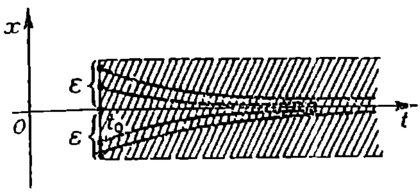


Fig. 20.3

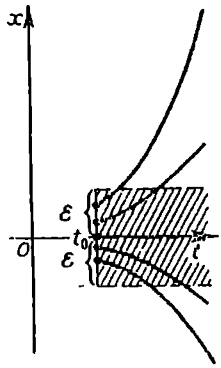


Fig. 20.4

Now consider a system of differential equations

$$\frac{dx_i}{dt} = f_i(t, x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n), \tag{20.8}$$

where f_i are defined for $a < t < +\infty$ and $x_1, x_2, \ldots, x_n$ ($x_1, x_2, \ldots, x_n$ lying in a domain D) and meet the conditions of the theorem on existence and uniqueness of the solution to the Cauchy problem. Suppose that all the solutions of (20.8) are extended without limit to the right for $t \geqslant t_0 > a$.

Definition. A solution $\varphi_i(t)$, $i = 1, 2, \ldots, n$, of the system (20.8) is said to be stable *in the sense of Laypunov* as $t \to +\infty$ if for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for any solution $x_i(t)$, $i = 1, 2, \ldots, n$, of the system, whose initial values obey

$$|x_i(t_0) - \varphi_i(t_0)| < \delta \qquad (i = 1, 2, \ldots, n)$$

the inequalities

$$|x_i(t) - \varphi_i(t)| < \varepsilon \qquad (i = 1, 2, \ldots, n) \tag{20.9}$$

are valid for all $t \geqslant t_0$, i.e., solutions with close initial values remain close for all $t \geqslant t_0$:

If for arbitrarily small $\delta > 0$ *at least for one* $x_i(t)$, $i = 1, 2, \ldots, n$, the inequalities (20.9) do not hold, then the solution $\varphi_i(t)$ is called *unstable*.

Definition. A solution $\varphi_i(t)$, $i = 1, 2, \ldots, n$, of (20.8) is called *asymptotically stable*, if

(1) the solution is stable;

(2) there exists $\delta_1 > 0$ such that any solution $x_i(t)$, $i = 1, 2, \ldots, n$, of the system, for which

$$|x_i(t_0) - \varphi_i(t_0)| < \delta_1 \qquad (i = 1, 2, \ldots, n)$$

satisfies the condition

$$\lim_{t \to +\infty} |x_i(t) - \varphi_i(t)| = 0 \qquad (i = 1, 2, \ldots, n).$$

Example. Using the definition of stability in the sense of Lyapunov show that the solution of the system

$$\frac{dx}{dt} = y, \qquad \frac{dx}{dt} = -x, \tag{*}$$

subject to the initial conditions

$$x(0) = 0, \qquad y(0) = 0, \tag{**}$$

is stable.

◀ A solution of (*) that meets the initial conditions (**) is $x(t) \equiv 0$, $y(t) \equiv 0$. A solution of the system meeting the conditions $x(0) = x_0$, $y(0) = y_0$ has the form

$$x(t) = x_0 \cos t + y_0 \sin t, \qquad y(t) = -x_0 \sin t + y_0 \cos t.$$

Take an arbitrary $\varepsilon > 0$ and show that there exists $\delta(\varepsilon) > 0$ such that for $|x_0 - 0| < \delta$ and $|y_0 - 0| < \delta$ we have

$$|x(t) - 0| = |x_0 \cos t + y_0 \sin t| < \varepsilon,$$
$$|y(t) - 0| = |-x_0 \sin t + y_0 \cos t| < \varepsilon$$

for all $t \geqslant 0$. This will exactly mean that, by definition, the solution $x(t) \equiv 0$, $y(t) \equiv 0$ of system (*) is stable in the sense of Lyapunov. Obviuosly, we have

$$|x_0 \cos t + y_0 \sin t| \leqslant |x_0 \cos t| + |y_0 \sin t| \leqslant |x_0| + |y_0|,$$
$$|-x_0 \sin t + y_0 \cos t| \leqslant |x_0 \sin t| + |y_0 \cos t| \leqslant |x_0| + |y_0|.$$

If we take $\delta(\varepsilon) = \varepsilon/2$, then for $|x_0| < \delta$ and $y_0 < \delta$ we will have

$$|x_0 \cos t + y_0 \sin t| < \varepsilon, \qquad |-x_0 \sin t + y_0 \cos t| < \varepsilon$$

for all $t \geqslant 0$, i.e., the zero solution will really be stable in the sense of Lyapunov although this stability will not be asymptotic. ▶

The fact that a nontrivial solution of a differential equation is stable *does not suggest* that the solution is bounded. Consider, say, the equation $dx/dt = 1$. The solution subject to the condition $x(0) = 0$ is the function $\varphi(t) = t$. The solution with the initial conditions $x(0) = x_0$ has the form $x(t) = t + x_0$. Geometrically, it is clear (Fig. 20.5) that for any $\varepsilon > 0$ there exists $\delta > 0$, e.g., $\delta = \varepsilon$, such that any solution $x(t)$ of the equation for which there holds the inequality $|x_0 - 0| < \delta$ meets the condition $|x(t) - t| < \varepsilon \ \forall t \geqslant 0$. This means that the solution $\varphi(t) = t$ is Lyapunov stable; this solution, however, is unbounded as $t \to +\infty$.

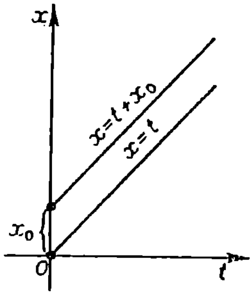


Fig. 20.5

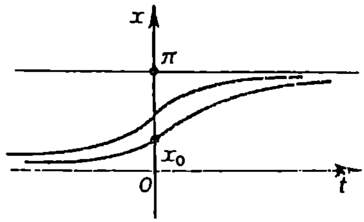


Fig. 20.6

Again the fact that solutions of an equation are bounded *does not suggest* that they are stable. Consider the equation

$$\frac{dx}{dt} = \sin^2 x. \tag{20.10}$$

Its obvious solutions are

$$x = k\pi \qquad (k = 0, \pm 1, \pm 2, \ldots). \tag{20.11}$$

Integrating (20.10) gives $\cot x = \cot x_0 - t$ or

$$x = \cot^{-1}(\cot x_0 - t), \qquad x_0 \neq k\pi. \tag{20.12}$$

All the solutions (20.11) and (20.12) are bounded in $(-\infty, +\infty)$. The solution $\varphi(t) \equiv 0$ is however unstable as $t \to +\infty$, since for any $x_0 \in (0, \pi)$ we have $\lim\limits_{t \to +\infty} x(t) = \pi$ (Fig. 20.6).

Boundedness and stability of solutions are thus independent concepts.

Remark. The solution $\varphi_i(t)$, $i = 1, 2, \ldots, n$, of (20.8) can always be transformed into the trivial solution $y_i \equiv 0$ of another system by the change $y_i = x_i(t) - \varphi_i(t)$, $i = 1, 2, \ldots, n$. Suppose that for simplicity we have one differential equation

$$\frac{dx}{dt} = f(t, x), \tag{*}$$

and we would like to examine for stability some solution $\varphi(t)$ of the equation.

We set $y(t) = x(t) - \varphi(t)$ (the quantity $x(t) - \varphi(t)$ is called a *perturbation*). Then $x(t) = y(t) + \varphi(t)$ and substitution into (*) yields

$$\frac{dy}{dt} + \frac{d\varphi}{dt} = f(t, y(t) + \varphi(t)). \tag{**}$$

But $\varphi(t)$ is a solution of (*), therefore $d\varphi/dt \equiv f(t, \varphi(t))$ and from (**) we have

$$\frac{dy}{dt} = f(t, y(t) + \varphi(t)) - f(t, \varphi(t)).$$

Denoting the right-hand side by $F(t, y)$ gives

$$\frac{dy}{dt} = F(t, y). \tag{***}$$

This equation has the solution $y \equiv 0$, since at $y \equiv 0$ its either side is identically zero in t:

$$F(t, 0) = f(t, \varphi(t)) - f(t, \varphi(t)) \equiv 0.$$

To summarize, the question of stability of the solution $\varphi(t)$ of (*) comes down to the question of stability of the trivial solution $y \equiv 0$ of the equation (***) to which (*) reduces. Therefore, in what follows we will as a rule assume that it is the trivial solution that is examined for stability.

20.3 Stability of Autonomous Systems. Simplest Types of Stationary Points

A normal system of differential equations is said to be *autonomous* if its right sides f_i are not explicity dependent on t, i.e., if it has the form

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n).$$

This means that the law of variation of unknown functions described by an autonomous system is not time-dependent, as it is normally the case with physical laws.

Consider an autonomous system

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n) \tag{20.13}$$

and let a collection of numbers $(a_1, a_2, \ldots, a_n)$ be such that

$$f_i(a_1, a_2, \ldots, a_n) = 0 \qquad (i = 1, 2, \ldots, n).$$

A system of functions $x_i(t) \equiv a_i$, $i = 1, 2, \ldots, n$, will then be a solution to (20.13). Point $(a_1, a_2, \ldots, a_n)$ in the phase space $(x_1, x_2, \ldots, x_n)$ is then called a *stationary* or *rest point* (*equilibrium*) of the system.

Let us take a system (20.13) for which $f_i(0, 0, \ldots, 0) = 0$, $i = 1, 2, \ldots, n$, so that $x_i = 0$, $i = 1, 2, \ldots, n$, is a stationary point of the system.

We denote by $S(R)$ the ball $\sum_{i=1}^{n} x_i^2 < R^2$. We then suppose that the system in $S(R)$ meets the conditions of the theorem on the existence and uniqueness of the solution of the Cauchy problem.

Definition. We will say that a stationary point $x_i = 0$, $i = 1, 2, \ldots, n$, of system (20.13) is *stable*, if for any $\varepsilon > 0$ $(0 < \varepsilon < R)$ there exists $\delta = \delta(\varepsilon) > 0$ such that any path of the system originating at the initial moment $t = t_0$ at a point $M_0 \in S(\delta)$ will then always stay within the confines of the ball $S(\varepsilon)$.

The stationary point is *asymptotically stable*, if:

(1) it is stable;

(2) there exists $\delta_1 > 0$ such that each path of the system originating at a point M_0 in $S(\delta_1)$ tends to the origin of coordinates, when the time t grows indefinitely (Fig. 20.7).

We will illustrate the definition.

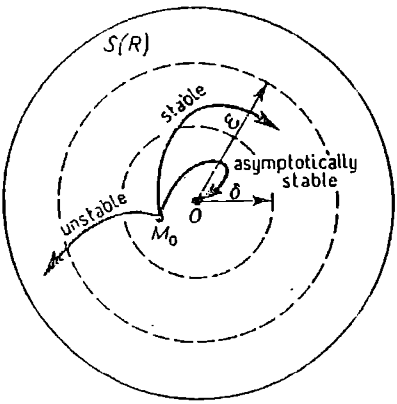


Fig. 20.7

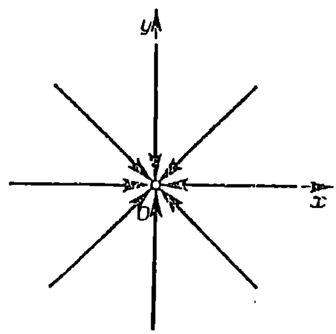


Fig. 20.8

Examples. (1) Consider the system

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = -x.$$

◄ The paths here are concentric circles $x^2 + y^2 = h^2$ with centre at the origin, which is the only stationary point of the system. If we take $\delta = \varepsilon$,

then any path originating within the circle $S(\delta)$ always remains within $S(\delta)$, and hence within $S(\varepsilon)$ as well, so that the situation is stable. But the paths do not approach the origin of coordinates as $t \to +\infty$ and the stationary point is not asymptotically stable. ►

(2) Consider the system

$$\frac{dx}{dt} = -x, \qquad \frac{dy}{dt} = -y.$$

◄ Its solution is $x = Ae^{-t}$, $y = Be^{-t}$. Hence $y/x = B/A = k = \text{const}$, therefore the paths are rays terminating at the origin (Fig. 20.8). We can again select $\delta = \varepsilon$. Any point on the path that at the initial moment of time lies within $S(\delta)$ always remains within the circle $S(\varepsilon)$, and also approaches indefinitely the origin as $t \to +\infty$. Consequently, the situation is asymptotically stable. ►

(3) Lastly, we take the system

$$\frac{dx}{dt} = x, \qquad \frac{dy}{dt} = y.$$

◄ Its solution is $x = Ae^{t}$, $y = Be^{t}$. Here also $y/x = k$ and the paths are the rays emerging from the origin of coordinates, but, unlike Example 2, the motion along the rays occurs away from the centre. The stationary point is unstable. ►

Simplest types of stationary points. We now examine the paths in the neighbourhood of the stationary point $x = 0$, $y = 0$ of the system of two linear homogeneous equations with constant coefficients

$$\frac{dx}{dt} = a_{11}x + a_{12}y, \qquad \frac{dy}{dt} = a_{21}x + a_{22}y, \tag{20.14}$$

where $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$. We will seek the solution in the form $x = \alpha e^{\lambda t}$, $y = \beta e^{\lambda t}$. We will find λ from the characteristic equation

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0. \tag{20.15}$$

Quantities α, β are found up to a factor from the system

$$(a_{11} - \lambda)\alpha + a_{12}\beta = 0, \qquad a_{21}\alpha + (a_{22} - \lambda)\beta = 0.$$

The following cases are possible:

(a) The roots λ_1, λ_2 of the characteristic equation (20.15) are *real* and *different*. The general solution of the system (20.14) has the form

$$\begin{aligned} x(t) &= C_1\alpha_1 e^{\lambda_1 t} + C_2\alpha_2 e^{\lambda_2 t}, \\ y(t) &= C_1\beta_1 e^{\lambda_1 t} + C_2\beta_2 e^{\lambda_2 t}. \end{aligned} \tag{20.16}$$

(1) Let $\lambda_1 < 0$, $\lambda_2 < 0$. The stationary point (0, 0) is then asymptotically stable, since the equation includes the factors $e^{\lambda_1 t}$, $e^{\lambda_2 t}$, and so all the points on any path that at $t = t_0$ are located within any δ-neighbourhood of the origin at sufficiently large t go into points lying within an arbitrary small ε-neighbourhood of the origin, and as $t \to +\infty$ they tend to the origin of coordinates. Such a stationary point is called a *stable node.*

At $C_2 = 0$ from (20.16) we get

$$x = C_1\alpha_1 e^{\lambda_1 t}, \qquad y = C_1\beta_1 e^{\lambda_1 t}.$$

It follows that $y = \beta_1 x/\alpha_1$ and the paths are two rays that enter the origin of coordinates with the slope $k_1 = \beta_1/\alpha_1$.

Similarly, at $C_1 = 0$ we also obtain two rays that enter the origin with the slope $k_2 = \beta_2/\alpha_2$.

Let now $C_1 \neq 0$ and $C_2 \neq 0$, and, for definiteness, let $|\lambda_1| > |\lambda_2|$. Then by (20.16)

$$\frac{dy}{dt} = \frac{C_1\beta_1\lambda_1 e^{\lambda_1 t} + C_2\beta_2\lambda_2 e^{\lambda_2 t}}{C_1\alpha_1\lambda_1 e^{\lambda_1 t} + C_2\alpha_2\lambda_2 e^{\lambda_2 t}} \to \frac{\beta_2}{\alpha_2} \quad \text{as} \quad t \to +\infty,$$

i.e., all the paths (except for the rays $y = \beta_1 x/\alpha_1$) in the vicinity of the stationary point $O(0, 0)$ have the same direction $y = \beta_2 x/\alpha_2$ (Fig. 20.9).

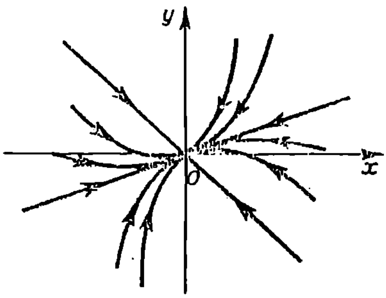


Fig. 20.9

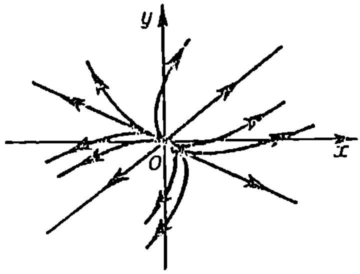


Fig. 20.10

(2) If $\lambda_1 > 0$, $\lambda_2 > 0$, then the paths are located as in the previous case, but the points move along them in the opposite direction. The stationary point of this type is called an *unstable node* (Fig. 20.10).

By way of example we consider the system

$$\frac{dx}{dt} = x, \qquad \frac{dy}{dt} = 2y.$$

For it $O(0, 0)$ is a stationary point. The characteristic equation

$$\begin{vmatrix} 1-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = 0$$

has the roots $\lambda_1 = 1$, $\lambda_2 = 2$, since the node is definitely unstable. We now pass from this system to one equation $dy/dx = 2y/x$ or $x\,dy - 2y\,dx = 0$. It has the solutions $y \equiv 0$, $x \equiv 0$ and $y = Cx^2$, since the system's paths will be rays coincident with the coordinate axes, and the family of parabolas tangent to the x-axis at the origin of coordinates (Fig. 20.11).

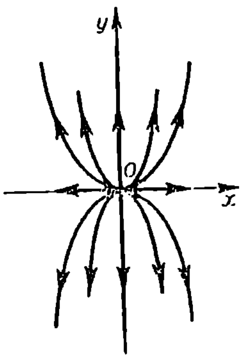


Fig. 20.11

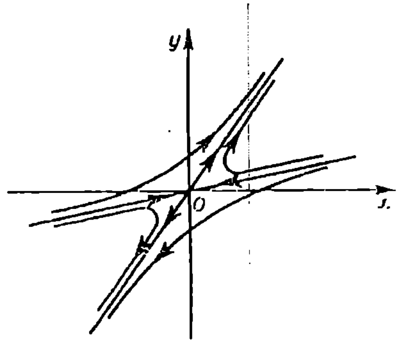


Fig. 20.12

(3) Let now $\lambda_1 > 0$, $\lambda_2 < 0$ (or we could also use $\lambda_1 < 0$, $\lambda_2 > 0$). Then the stationary point is unstable. At $C_2 = 0$ we obtain the motion

$$x = C_1\alpha_1 e^{\lambda_1 t}, \qquad y = C_1\beta_1 e^{\lambda_1 t}$$

such that the point will move with t along the ray $y = \beta_1 x/\alpha_1$ in the direction away from the origin ($\lambda_1 > 0$) receding from it indefinitely. At $C_1 = 0$ we have

$$x = C_2\alpha_2 e^{\lambda_2 t}, \qquad y = C_2\beta_2 e^{\lambda_2 t}.$$

It follows that as t increases the point moves along the ray $y = \beta_2 x/\alpha_2$ in the direction to the origin ($\lambda_2 < 0$). If $C_1 \neq 0$ and $C_2 \neq 0$, then as $t \to +\infty$ and as $t \to -\infty$ the path leaves the vicinity of the stationary point. A stationary point of this type is known as a *saddle point* (Fig. 20.12).

Examine the nature of the stationary point $O(0, 0)$ of the system

$$\frac{dx}{dt} = -x, \qquad \frac{dy}{dt} = y. \tag{20.17}$$

The characteristic equation of the system $\lambda^2 - 1 = 0$ has the roots $\lambda_1 = 1$, $\lambda_2 = -1$. We now pass on to one equation

$$\frac{dy}{dx} = -\frac{y}{x} \quad \text{or} \quad x\,dy + y\,dx = 0, \tag{20.18}$$

integrating which we obtain $xy = C$.

Other solutions of (20.18) are $y \equiv 0$ and $x \equiv 0$.

In consequence, the integral curves of this equation (or the paths of system (20.17)) are equilateral hyperbolas and rays coincident with the coordinate semiaxes.

(b) The roots λ_1, λ_2 of the characteristic equation are *complex*: $\lambda_{1,2} = p \pm iq$, $q \neq 0$. The general solution of system (20.14) can be represented in the form

$$\begin{aligned} x(t) &= e^{pt}(C_1 \cos qt + C_2 \sin qt), \\ y(t) &= e^{pt}(C_1^* \cos qt + C_2^* \sin qt), \end{aligned} \tag{20.19}$$

where C_1 and C_2 are arbitrary constants, and C_1^*, C_2^* are some linear combinations of these constants.

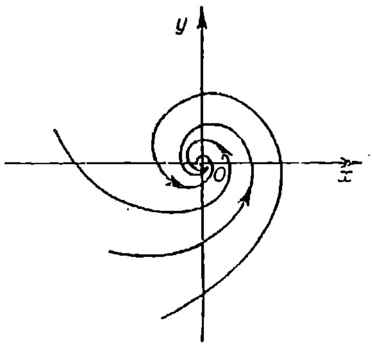


Fig. 20.13

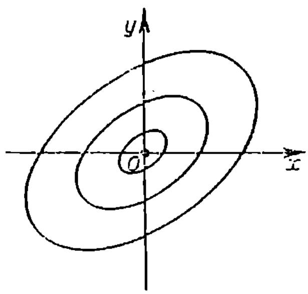


Fig. 20.14

(1) Let $\lambda_{1,2} = p + iq$, $p < 0$, $q \neq 0$. The factor e^{pt}, $p < 0$, will then tend to zero as $t \to +\infty$, and the second factors in (20.19) are bounded periodic functions. The paths will be spirals that asymptotically tend to the origin of coordinates when $t \to +\infty$. The stationary point $x = 0$, $y = 0$ is asymptotically stable. It is called a *stable focus* (Fig. 20.13).

(2) If $\lambda_{1,2} = p \pm iq$, $p > 0$, $q \neq 0$, then substituting $-t$ for t brings us back to the previous case. The paths are the same as previously, but the motion along them with increasing t occurs in the opposite direction. The stationary point is unstable and is called an *unstable focus*.

(3) If then $\lambda_{1,2} = \pm iq$, $q \neq 0$, then the solutions to (20.14) are periodic functions. The paths are closed curves that contain a stationary point, which is then called the *centre* (Fig. 20.14). A centre is a stable stationary point; however, there is no asymptotic stability, since the solutions

$$x(t) = C_1 \cos qt + C_2 \sin qt, \qquad y(t) = C_1^* \cos qt + C_2^* \sin qt$$

do not tend to zero when $t \to +\infty$.

By way of example we consider the system

$$\frac{dx}{dt} = ax - y, \qquad \frac{dy}{dt} = x + ay \qquad (a = \text{const}). \tag{20.20}$$

The characteristic equation of the system $(a - \lambda)^2 + 1 = 0$ has the complex roots $\lambda_{1,2} = a \pm i$.

We pass from the system to the equation

$$\frac{dy}{dx} = \frac{x + ay}{ax - y} \tag{20.21}$$

and introduce the polar coordinates $x = \varrho \cos \varphi$, $y = \varrho \sin \varphi$, then $\varrho^2 = x^2 + y^2$, $\tan \varphi = y/x$ and

$$\varrho \frac{d\varrho}{dx} = x + y \frac{dy}{dx}, \qquad \varrho^2 \frac{d\varphi}{dx} = x \frac{dy}{dx} - y$$

and hence

$$\frac{d\varrho}{d\varphi} = \varrho \frac{x + yy'}{xy' - y}.$$

Using equation (20.21) we find that

$$\frac{d\varrho}{d\varphi} = a\varrho, \qquad \text{hence} \qquad \varrho = Ce^{a\varphi}.$$

These integral curves are logarithmic spirals centred around the origin of coordinates, which is reached in the limit as $\varphi \to +\infty$ or $\varphi \to -\infty$ depending on whether or not $a < 0$ or $a > 0$. We thus have a stationary point of focus type. In the special case of $a = 0$ equation (20.21) becomes

$$\frac{dy}{dx} = -\frac{x}{y}.$$

The integral curves of this equation are circles with centre at the origin, which at $a = 0$ is a centre-type stationary point of system (20.20).

(c) Roots λ_1, λ_2 of the characteristic equation are *multiple*, i.e., $\lambda_1 = \lambda_2$. This case is quite rare. Even a negligible change in the coefficients of the system destroys it. Using, say, the elimination method, we find that the

general solution of the system (20.14) has the form

$$x(t) = (C_1 + C_2 t)e^{\lambda_1 t},$$
$$y(t) = (C_1^* + C_2^* t)e^{\lambda_1 t},$$

where C_1^*, C_2^* are some linear combinations of C_1 and C_2.

(1) If $\lambda_1 = \lambda_2 < 0$, then the solution contains the factor $e^{\lambda_1 t}$, $\lambda_1 < 0$, and so $x(t)$, $y(t)$ tend to zero as $t \to +\infty$. The stationary point $x = 0$, $y = 0$ is asymptotically stable. It is called a *stable degenerate node* (Fig. 20.15).

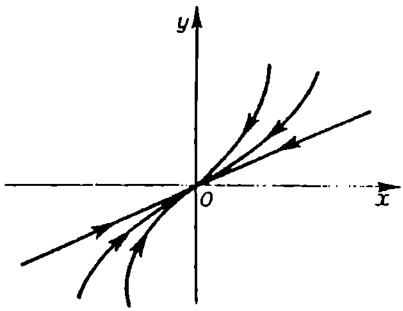


Fig. 20.15

It differs from a node of case (a1) by the fact that there one path had a tangent distinct from all the others. Also possible is a *dicritical node* (see Fig. 20.8).

(2) At $\lambda_1 = \lambda_2 > 0$ substitution of $-t$ for t brings us back to the previous case, but now the point moves along the paths in the opposite direction. The stationary point is now called *unstable degenerate node*.

For example, for the system

$$\frac{dx}{dt} = x, \qquad \frac{dy}{dt} = x + y$$

the characteristic equation is $(\lambda - 1)^2 = 0$. It has multiple roots $\lambda_1 = \lambda_2 = 1$. Dividing the first equation of the system by the second one gives

$$\frac{dy}{dx} = 1 + \frac{y}{x},$$

hence

$$y = x(\ln|x| + C).$$

In that case

$$\lim_{x\to 0} y(x) = 0, \qquad \lim_{x\to 0} y'(x) = \lim_{x\to 0} (1 + C + \ln|x|) = -\infty.$$

Therefore, all the integral curves pass through the origin of coordinates and for all of them the y-axis is a common tangent there.

We have exhausted all the possibilities, since the case of $\lambda_1 = 0$ (or $\lambda_2 = 0$) is excluded by the condition

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

Example. Examine the equation of small oscillations of a pendulum in the presence of friction.

◄ The equation in this case has the form

$$\frac{d^2x}{dt^2} = -x - k\frac{dx}{dt}, \tag{*}$$

where x is the angle of small displacement of the pendulum from the vertical, k is the friction coefficient. We replace equation (*) by the equivalent system

$$\frac{dx}{dt} = x_1, \qquad \frac{dx_1}{dt} = -x - kx_1. \tag{**}$$

The characteristic equation for (**) is

$$\begin{vmatrix} -\lambda & 1 \\ -1 & -k-\lambda \end{vmatrix} = 0 \qquad \text{or} \qquad \lambda^2 + k\lambda + 1 = 0.$$

It has the roots $\lambda_{1,2} = -k/2 \pm \sqrt{k^2/4 - 1}$. If $0 < k < 2$, then these roots will be complex numbers with a negative real part, so that the lower equilibrium position of the pendulum $x = x_1 = 0$ will be a stable focus. Equation (*) has the solution

$$x(t) = Ae^{-kt/2}\sin(\omega t + \alpha),$$

where $\omega = \sqrt{1 - k^2/4}$ is the frequency of oscillations, and the quantities A, α are determined from the initial conditions.

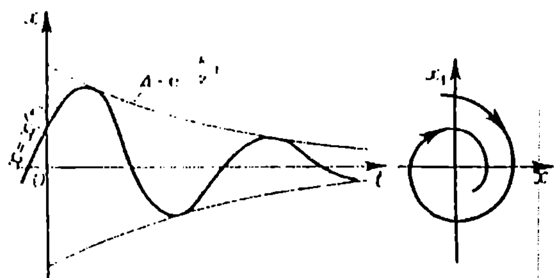


Fig. 20.16

The solution and the phase curve for $0 < k < 2$ have a form as shown in Fig. 20.16. As $k \to 0$, i.e., as the friction coefficient decreases, the focus turns into a centre: the pendulum will undergo undamped periodic oscillations. ►

We will now summarize the results concerning the stability of solutions of a system of n linear homogeneous differential equations of the first order with constant coefficients

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}x_j, \qquad i = 1, 2, \ldots, n, \qquad a_{ij} = \text{const.} \tag{20.22}$$

Let us examine the characteristic equation

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & & a_{1n} \\ a_{21} & a_{22} - \lambda & & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0$$

of system (20.22).

We can make the following statements:

(1) If all the roots of a characteristic equation have a negative real part, then all the solutions of (20.22) are asymptotically stable. It appears that all the terms of the general solution then contain the factors $e^{\mathrm{Re}\lambda_k t}$ that tend to zero as $t \to +\infty$;

(2) If at least one root λ_k of the characteristic equation has a positive real part, then the solutions of the system are all unstable;

(3) If the characteristic equation has *simple* roots with zero real parts (i.e., purely imaginary or zero roots), and other roots, if any, have negative real parts, then the solutions are all stable, although not asymptotically so.

These results also apply to one linear differential equation with constant coefficients.

It is worth noting that for a linear system all the solutions are either stable or unstable at the same time.

Theorem 20.2. *Solutions of the system of linear differential equations*

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}(t)x_j + f_i(t) \qquad (i = 1, 2, \ldots, n), \tag{20.23}$$

are all and at the same time either stable or unstable.

◄ We transform any particular solution $\varphi_i(t)$, $i = 1, 2, \ldots, n$, of system (20.23) into a trivial one by a change $y_i = x_i(t) - \varphi_i(t)$. System (20.23) then yields a linear homogeneous system in $y_i(t)$

$$\frac{dy_i}{dt} = \sum_{j=1}^{n} a_{ij}(t)y_j \qquad (i = 1, 2, \ldots, n). \tag{20.24}$$

Consequently, all the particular solutions of system (20.23) in terms of stability behave similarly, namely as a trivial solution of the homogeneous system (20.24).

Let the trivial solution $y_i(t) \equiv 0$, $i = 1, 2, \ldots, n$, of system (20.24) be stable. This means that for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for any other solution $y_i(t)$, $i = 1, 2, \ldots, n$, of the system it follows from the condition $|y_i(t_0)| < \delta$, $i = 1, 2, \ldots, n$, that

$$|y_i(t)| < \varepsilon \qquad (i = 1, 2, \ldots, n) \qquad \forall t \geqslant t_0.$$

Notice that $y_i(t) = x_i(t) - \varphi_i(t)$, and so we find that from the condition

$$|x_i(t_0) - \varphi_i(t_0)| < \delta \qquad (i = 1, 2, \ldots, n),$$

it follows that

$$|x_i(t) - \varphi_i(t)| < \varepsilon \qquad (i = 1, 2, \ldots, n) \qquad \forall t \geqslant t_0,$$

for any solution $x_i(t)$, $i = 1, 2, \ldots, n$ of the original system (20.23). By definition, this means that the solution $x_i(t)$, $i = 1, 2, \ldots, n$, of the system is stable. ►

This does not hold good for nonlinear systems. Some of their solutions may be stable and the others unstable.

Consider, for example, the nonlinear equation

$$\frac{dx}{dt} = 1 - x^2.$$

It has the obvious solution $x(t) = -1$ and $x(t) = 1$. The solution $x(t) = -1$ is unstable, and the solution $x(t) = 1$ is asymptotically stable. Indeed, when $t \to +\infty$ all the solutions

$$x(t) = \frac{(1 + x_0)e^{2(t - t_0)} - (1 - x_0)}{(1 + x_0)e^{2(t - t_0)} + (1 - x_0)}, \qquad x_0 \neq -1,$$

tend to $+1$. This means, by definition, that the solution $x(t) = 1$ is asymptotically stable.

Remark. In the case of $n = 3$, we can also examine the locations of the paths about the stationary point $O(0, 0, 0)$ of (20.22). Here the so-called node foci (Fig. 20.17), saddle foci (Fig. 20.18), and so on, are possible.

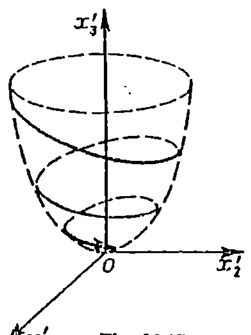


Fig. 20.17

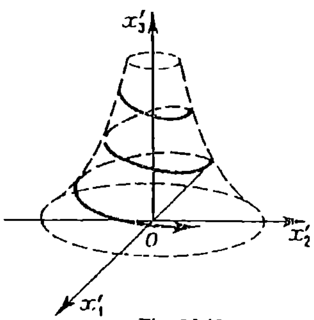


Fig. 20.18

20.4 Method of Lyapunov's Functions

The method examines the stability of a stationary point of a system of differential equations using an adequately chosen function $v(t, x_1, x_2, \ldots, x_n)$ — the so-called *Lyapunov's function* — and this operation is carried out without preliminarily solving the system. Here lies one significant advantage of the method.

We will confine our discussion to the autonomous system

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n), \tag{20.25}$$

for which $x_i = 0$, $i = 1, 2, \ldots, n$, is a stationary point.

The idea behind the method is as follows. Suppose that we examine for stability a stationary point $x_i = 0$, $i = 1, 2, \ldots, n$, of (20.25). If the points of all paths would approach the origin as t increases or, at least, would not recede from it, then the stationary point under consideration would be stable. The test for this condition does not require the knowledge of the solutions of the system. If ϱ is the distance of the point of the path $x_i = x_i(t)$, $i = 1, 2, \ldots, n$, to the origin of coordinates

$$\varrho = \sqrt{\sum_{i=1}^{n} x_i^2(t)},$$

then

$$\frac{d\varrho}{dt} = \sum_{i=1}^{n} \frac{\partial \varrho}{\partial x_i}\frac{dx_i}{dt} = \sum_{i=1}^{n} \frac{\partial \varrho}{\partial x_i} f_i(x_1, x_2, \ldots, x_n) \tag{20.26}$$

is the derivative along the path. The right-hand side of (20.26) is the known function of $x_1, x_2, \ldots, x_n$, and we can examine it for sign. If it appears that $d\varrho/dt \leqslant 0$, then the points on all paths do not recede from the origin as t increases and the stationary point $x_i = 0$, $i = 1, 2, \ldots, n$, is stable. A stationary point may be stable, however, even when approached by points of the paths in a nonmonotone manner with increasing t (e.g., when the paths are ellipses). Therefore, Lyapunov, instead of the function ϱ, considered the function $v(x_1, x_2, \ldots, x_n)$, which is in a way a "generalized distance" from the origin of coordinates.

Definition. A function $v(x_1, x_2, \ldots, x_n)$ defined in a neighbourhood of the origin of coordinates is said to be *definite* (*positive definite* or *negative definite*) if in a domain G

$$|x_i| \leqslant h \qquad (i = 1, 2, \ldots, n),$$

where h is a sufficiently small positive number, it may take on values of one sign only and becomes zero only at $x_1 = x_2 = \ldots = x_n = 0$.

So, in the case when $n = 3$ the functions

$$v = x_1^2 + x_2^2 + x_3^2 \qquad \text{and} \qquad v = x_1^2 + 2x_1x_2 + 2x_2^2 + x_3^2$$

will be positive definite. Here the quantity $h > 0$ may be taken arbitrarily large.

A function $v(x_1, x_2, \ldots, x_n)$ is said to be of *constant sign* (*positive* or *negative*), if in a region G it may take on values of one sign only, but can also become zero at $x_1^2 + x_2^2 + \ldots + x_n^2 \neq 0$.

For example, the function

$$v(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + x_2^2 + x_3^2$$

is of constant sign (positive). It is easily seen that we can represent $v(x_1, x_2, x_3)$ as $v(x_1, x_2, x_3) = (x_1 + x_2)^2 + x_3^2$. It follows that it is nonnegative everywhere, but vanishes at $x_1^2 + x_2^2 + x_3^2 \neq 0$, namely, at $x_3 = 0$ and any x_1, x_2, such that $x_1 = -x_2$.

Let $v(x_1, x_2, \ldots, x_n)$ be a differentiable function and let $x_1(t)$, $x_2(t), \ldots, x_n(t)$ be some functions of time that obey the system of differential equations (20.25). Then the total derivative of v with respect to time will be

$$\frac{dv}{dt} = \sum_{i=1}^{n} \frac{\partial v}{\partial x_i}\frac{dx_i}{dt} = \sum_{i=1}^{n} \frac{\partial v}{\partial x_i} f_i(x_1, x_2, \ldots, x_n). \tag{20.27}$$

The quantity dv/dt given by (20.27) is called the *total derivative* of v with respect to time derived by virtue of (20.25).

A *Lyapunov function* is a function $v(x_1, x_2, \ldots, x_n)$ that possesses the following properties:

(1) it is differentiable in a certain neighbourhood Ω of the origin of coordinates;

(2) it is positive (negative) definite in Ω and $v(0, 0, \ldots, 0) = 0$;

(3) its total derivative dv/dt derived by virtue of (20.25) is subject to

$$\frac{dv}{dt} \leqslant 0 \qquad \left(\frac{dv}{dt} \geqslant 0\right)$$

everywhere in Ω.

Theorem 20.3 (Lyapunov's theorem on stability). *If for a system of differential equations*

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n),$$
$$f_i(0, 0, \ldots, 0) = 0, \tag{20.28}$$

there exists a definite differentiable function $v(x_1, x_2, \ldots, x_n)$, whose total derivative dv/dt with respect to time derived by virtue of (20.28) is a func-

tion of constant sign opposite to v, or it is identically zero, then the stationary point $x_i = 0$ $(i = 1, 2, \ldots, n)$ of (20.28) *is stable.*

We will just sketch the idea of the proof. Suppose for definiteness that $v(x_1, x_2, \ldots, x_n)$ is a positive definite function for which $dv/dt \leqslant 0$. Since $v(x_1, x_2, \ldots, x_n) \geqslant 0$, and $v = 0$ only at $x_1 = x_2 = \ldots = x_n = 0$, then the origin of coordinates is a point where $v(x_1, x_2, \ldots, x_n)$ has a strict minimum. In the neighbourhood of the origin the surfaces of the level $v(x_1, x_2, \ldots, x_n) = C$ of the function v can be shown to be closed surfaces that enclose the origin. To clarify the picture we will take the case of $n = 2$. Since $v \geqslant 0$ for small x_1, x_2 and $v = 0$ only for $x_1 = x_2 = 0$, then the surface $z = v(x_1, x_2)$ looks rather like a paraboloid as shown in Fig. 20.19.

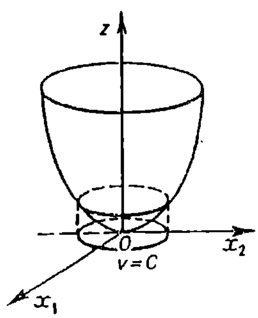


Fig. 20.19

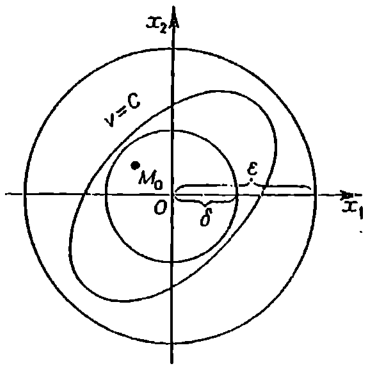


Fig. 20.20

The lines of the level $v(x_1, x_2) = C$ are a family of closed curves around the origin of coordinates. If $C_1 < C_2$, then the lines $v = C_1$ wholly lie within the line $v = C_2$. We put $\varepsilon > 0$. At sufficiently small $C > 0$ the line $v = C$ wholly belongs to the ε-neighbourhood of the origin, although it never passes through it. We can thus choose $\delta > 0$ such that the δ-neighbourhood of the origin will wholly lie within the line $v = C$. In that neighbourhood $v < C$ (Fig. 20.20).

Consider the path of (20.28) that at $t = t_0$ originates at a point $M_0(x_1(t_0), x_2(t_0))$ in the δ-neighbourhood of the origin of coordinates. This path will never cross any of the lines $v(x_1, x_2) = C$ from inside with increasing t. If such a crossing were possible at some point, then at that point or in its neighbourhood the function $v(x_1(t), x_2(t))$ would of necessity have a positive derivative dv/dt, since in passing from some line $v = C$ to another line of the family that encompasses the first one, the function increases. But this is impossible since as stated $dv/dt \leqslant 0$. Therefore, if at the initial moment some path was inside the region bounded by the line

$v = C$, then it will for ever remain within this region. It is clear thus that for any $\varepsilon > 0$ there exists $\delta > 0$ such that any path of the system that at $t = t_0$ originated from an ε-neighbourhood of the origin of coordinates will for all $t \geqslant t_0$ belong to the ε-neighbourhood of the origin. This exactly means that the stationary point $x_i = 0$, $i = 1, 2, \ldots, n$, of (20.28) is stable.

Theorem 20.4 (Lyapunov's theorem on asymptotic stability). *If for the system*

$$dx_i/dt = f_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n), \tag{20.29}$$

there exists a definite differentiable function $v(x_1, x_2, \ldots, x_n)$, such that its total derivative with respect to time derived by virtue of the system, is also a definite function of a sign opposite to v, then the stationary point $x_i = 0$, $i = 1, 2, \ldots, n$, of (20.28) is asymptotically stable.

Examples. (1) Examine for stability the stationary point $O(0, 0)$ of the system

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = -x. \tag{*}$$

◄ Suppose that $v(x, y) = x^2 + y^2$. This function is positive definite. By virtue of (*) we find

$$\frac{dv}{dt} = \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt} = 2xy - 2xy \equiv 0.$$

It follows from Theorem 20.3 that the stationary point $O(0, 0)$ of system (*) is stable (centre). There is no asymptotic stability: the paths of (*) are circles, and they do not tend to $O(0, 0)$ as $t \to +\infty$. ►

(2) Examine for stability the stationary point $O(0, 0)$ of the system

$$\frac{dx}{dt} = y - x^3, \qquad \frac{dy}{dt} = -x - y^3. \tag{**}$$

◄ If again we put $v(x, y) = x^2 + y^2$, we will find

$$\frac{dv}{dt} = 2x(y - x^3) + 2y(-x - y^3) = -2(x^4 + y^4).$$

Therefore, dv/dt is a negative definite function. By Theorem 20.4 the stationary point $O(0, 0)$ of system (**) is asymptotically stable.

Theorem 20.5 (on instability). *Suppose that for the system of differential equations*

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n)\ (f_i(0, 0, \ldots, 0) = 0),\ i = 1, 2, \ldots, n, \tag{20.30}$$

there exists a function $v(x_1, x_2, \ldots, x_n)$ differentiable about the origin of coordinates, such that $v(0, 0, \ldots, 0) = 0$. If its total derivative dv/dt der-

ived by virtue of system (20.30) *is a positive definite function and at an arbitrarily small distance from the origin there are points where* $v(x_1, x_2, \ldots, x_n)$ *assumes positive values, then the stationary point* $x_i = 0$, $i = 1, 2, \ldots, n$, *of system* (20.30) *is unstable.*

Example. Examine for stability the stationary point $O(0, 0)$ of the system

$$\frac{dx}{dt} = x, \qquad \frac{dy}{dt} = -y.$$

◄ We take the function $v(x, y) = x^2 - y^2$. The function

$$\frac{dv}{dt} = \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt} = 2(x^2 + y^2)$$

will then be positive definite. Since at an arbitrarily small distance from the origin there are points for which $v > 0$ (e.g., $v = x^2 > 0$ along the straight line $y = 0$), then the conditions of Theorem 20.5 are met and the stationary point $O(0, 0)$ is unstable (saddle). ►

The method of Lyapunov's functions is universal and effective for a wide variety of problems in stability theory.

The method has the disadvantage that there does not exist a sufficiently general procedure of constructing Lyapunov's functions.

In the simplest cases a Lyapunov function can be sought for in the form

$$v(x, y) = ax^2 + by^2, \; v(x, y) = ax^4 + by^4, \; a > 0, \; b > 0 \text{ and so on.}$$

20.5 Stability in First (Linear) Approximation

Let

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, 2, \ldots, n) \tag{20.31}$$

and let $x_i = 0$, $i = 1, 2, \ldots, n$, be a stationary point of the system, i.e.,

$$f_i(0, 0, \ldots, 0) = 0 \qquad (i = 1, 2, \ldots, n). \tag{20.32}$$

Suppose that $f_i(x_1, x_2, \ldots, x_n)$ are differentiable in a neighbourhood of the origin of coordinates sufficient number of times. Using the Taylor formula we expand f_i in x about the origin:

$$f_i(x_1, x_2, \ldots, x_n) = f_i(0, 0, \ldots, 0) + \sum_{j=1}^{n} \frac{\partial f_i(0, 0, \ldots, 0)}{\partial x_j} x_j + R_i(x_1, x_2, \ldots, x_n)$$

or, by (20.32),

$$f_i(x_1, x_2, \ldots, x_n) = \sum_{j=1}^{n} a_{ij}x_j + R_i(x_1, x_2, \ldots, x_n),$$

where $a_{ij} = \partial f_i(0, 0, \ldots, 0)/\partial x_j = \text{const}$, and R_i contain terms of no less than the second order of smallness in $x_1, x_2, \ldots, x_n$. The system of differential equations (20.31) becomes

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}x_j + R_i(x_1, x_2, \ldots, x_n),$$
$$i = 1, 2, \ldots, n, \qquad a_{ij} = \text{const}. \tag{20.33}$$

Since the concept of stability of the stationary point $O(0, 0, \ldots, 0)$ is associated with a small neighbourhood of the origin of coordinates in phase space, it would be natural to expect that the behaviour of solutions of (20.31) will be governed by the main linear terms of the expansion of f_i in x. Therefore, along with (20.33) we will consider the system

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij}x_j \qquad (i = 1, 2, \ldots, n), \tag{20.34}$$

referred to as a *system of equations in the first (linear) approximation* for system (20.33).

Generally speaking, there is no strict connection between (20.33) and (20.34). Consider for example

$$\frac{dx}{dt} = x^2. \tag{20.35}$$

Here $f(x) \equiv x^2$; the linearized equation for (20.35) takes the form

$$\frac{dx}{dt} = 0. \tag{20.36}$$

The solution $x(t) \equiv 0$ of (20.36) is stable. At the same time, being a solution of the original equation (20.35), it is unstable for it. It is easily seen that each real solution of (20.35) that meets the initial conditions $x|_{t=0} = x_0 > 0$ has the form $x = x_0/(1 - tx_0)$ and it is nonexistent at $t = 1/x_0$ (the solution is not extendible to the right).

Theorem 20.6. *If all the roots of the characteristic equation*

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & & a_{1n} \\ a_{21} & a_{22} - \lambda & & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{20.37}$$

have negative real parts, then the stationary point $x_i = 0$, $i = 1, 2, \ldots, n$, *of* (20.34) *and* (20.33) *is asymptotically stable.*

If the conditions of the theorem are met, an examination for stability in the first approximation is possible.

Theorem 20.7. *If at least one root of the characteristic equation* (20.37) *has a positive real part, then the stationary point* $x_i = 0$ *of* (20.34) *and* (20.33) *is unstable.*

In that case it is also possible to examine it for stability in the first approximation.

We will just give an outline of the proof of Theorems 20.6 and 20.7. Suppose for simplicity that the roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the characteristic equation (20.36) are real and different. There exists then a nondegenerate matrix $\mathbf{T}$ with constant elements, such that the matrix $\mathbf{T}^{-1}\mathbf{AT}$ will be diagonal:

$$\mathbf{T}^{-1}\mathbf{AT} = \begin{pmatrix} \lambda_1 & 0 & 0 & & 0 \\ 0 & \lambda_2 & 0 & & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

where $\mathbf{A} = (a_{ij})_{i,j=1}^{n}$ is a matrix of the coefficients of system (20.34). We put

$$\mathbf{X} = \mathbf{TY}, \qquad \text{where } \mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Then $\dfrac{d\mathbf{X}}{dt} = \mathbf{T}\dfrac{d\mathbf{Y}}{dt}$, and system (20.34) becomes $\mathbf{T}\dfrac{d\mathbf{Y}}{dt} = \mathbf{ATY}$. We obtain $d\mathbf{Y}/dt = \mathbf{T}^{-1}\mathbf{ATY}$, or due to the choice of matrix $\mathbf{T}$

$$\frac{dy_i}{dt} = \lambda_i y_i \qquad (i = 1, 2, \ldots, n).$$

System (20.33), if subjected to the same transformations, will yield

$$\frac{dy_i}{dt} = \lambda_i y_i + \bar{R}_i(y_1, y_2, \ldots, y_n). \tag{20.38}$$

Now again $\bar{R}_i$ will include the terms of no less than the second order of smallness in y_i as $y_i \to 0$.

Consider the following possibilities:

(1) All the roots λ_k are negative. Let

$$v = y_1^2 + y_2^2 + \ldots + y_n^2,$$

then the derivative dv/dt will, by (20.38), have the form

$$\frac{dv}{dt} = 2(\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2) + S(y_1, y_2, \ldots, y_n),$$

where $S(y_1, y_2, \ldots, y_n)$ as $\sum_{i=1}^{n} y_i^2 \to 0$ is a small quantity of a higher order than $\sum_{k=1}^{n} \lambda_k y_k^2$.

Thus in a sufficiently small neighbourhood Ω of a point $O(0, 0, \ldots, 0)$ the function $v(y_1, y_2, \ldots, y_n)$ is positive definite, and the derivative dv/dt is negative definite, and so the stationary point $O(0, 0, \ldots, 0)$ is asymptotically stable.

(2) Some of the roots λ_k (e.g., $\lambda_1, \lambda_2, \ldots, \lambda_m$, $m \leqslant n$) are positive, the others are negative. Suppose that

$$v = y_1^2 + y_2^2 + \ldots + y_m^2 - y_{m+1}^2 - \ldots - y_n^2,$$

then

$$\frac{dv}{dt} = 2(\lambda_1 y_1^2 + \ldots + \lambda_m y_m^2 - \lambda_{m+1} y_{m+1}^2 - \ldots - \lambda_n y_n^2) + S(y_1, y_2, \ldots, y_n).$$

It is seen that arbitrarily close to the origin there are points (e.g., such that $y_{m+1} = \ldots = y_n = 0$), where $v > 0$. Since $\lambda_{m+1}, \ldots, \lambda_n$ are negative, the derivative dv/dt is a positive definite function. By virtue of Theorem 20.5 the stationary point $O(0, 0, \ldots, 0)$ is unstable.

In the *critical* case where all the real parts of the roots of the characteristic equation are nonpositive, and the real part of at least one root is zero, the stability of the trivial solution of system (20.33) begins to be influenced by the nonlinear terms R_i, which makes impossible any examination for stability in the first approximation.

Examples. (1) Examine for stability in the first approximation the stationary point $x = 0$, $y = 0$ of the system

$$\frac{dx}{dt} = -x + 2y - 5y^2, \qquad \frac{dy}{dt} = 2x - y + \frac{x^3}{2}. \tag{*}$$

◄ A first-approximation system has the form

$$\frac{dx}{dt} = -x + 2y, \qquad \frac{dy}{dt} = 2x - y. \tag{**}$$

The nonlinear terms meet the adequate conditions: their order is not less than two. We set up the characteristic equation for (**)

$$\begin{vmatrix} -1-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix} = 0 \quad \text{or} \quad \lambda^2 + 2\lambda - 3 = 0.$$

The roots of the characteristic equations are $\lambda_1 = 1, \lambda_2 = -3$. Since $\lambda_1 > 0$, the zero solution $x \equiv 0$, $y \equiv 0$ of (*) is unstable. ▶

(2) Examine for stability the stationary point $O(0, 0)$ of the system

$$\frac{dx}{dt} = y - x^3, \qquad \frac{dy}{dt} = -x - y^3. \tag{*}$$

The stationary point $x = 0, y = 0$ of system (*) is asymptotically stable, since for this system Lyapunov's function $v = x^2 + y^2$ satisfies the conditions of the Lyapunov theorem on asymptotic stability. Specifically,

$$\frac{dv}{dt} = 2x(y - x^3) + 2y(-x - y^3) = -2(x^4 + y^4) \leqslant 0.$$

At the same time the stationary point $x = 0$, $y = 0$ of the system

$$\frac{dx}{dt} = y + x^3, \qquad \frac{dy}{dt} = -x + y^3 \tag{**}$$

is unstable.

For $v(x, y) = x^2 + y^2$ we by (**) have

$$\frac{dv}{dt} = \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt} = 2x(y + x^3) + 2y(-x + y^3) = 2(x^4 + y^4),$$

i.e., dv/dt is a positive definite function. At an arbitrarily small distance to the origin $O(0, 0)$ there are points where $v(x, y) > 0$.

By Theorem 20.5 we conclude that the stationary point $O(0, 0)$ of system (**) is unstable.

For systems (*) and (**) a first-approximation system will be the same

$$\frac{dx}{dt} = y, \qquad \frac{dy}{dt} = -x. \tag{***}$$

The characteristic equation $\lambda^2 + 1 = 0$ for (***) has purely imaginary roots — critical case (the real parts of the roots of the characteristic equation are zero). For the first-approximation system (***) the origin of coordinates is a stable stationary point (centre). Systems (*) and (**) are obtained by small perturbation of the right-hand sides of (***) in a neighbourhood of the origin of coordinates. But due to these small perturbations the stationary point $O(0, 0)$ for system (*) becomes asymptotically stable and for system (**) unstable.

This example shows that in a critical case nonlinear terms may influence the stability of a stationary point. ▶

Problem. Examine for stability the stationary point $O(0, 0)$ of the system

$$\frac{dx}{dt} = y - xf(x, y), \quad \frac{dy}{dt} = -x - yf(x, y),$$

where the function $f(x, y)$ is expandable into the convergent power series and $f(0, 0) = 0$.

Exercises

Using the definition, examine for stability the solutions of the following equations:

1. $\frac{dx}{dt} + x = 1$, $x(0) = 1$.
2. $\frac{dx}{dt} - x = 1$, $x(0) = -1$.
3. $\frac{dx}{dt} = 2$, $x(0) = 0$.

Establish the nature of the stationary point $O(0, 0)$ of the system and draw the paths in a neighbourhood of that point:

4. $\begin{cases} \frac{dx}{dt} = x - 2y, \\ \frac{dy}{dt} = 3x - 4y. \end{cases}$
5. $\begin{cases} \frac{dx}{dt} = x + 2y, \\ \frac{dy}{dt} = 2x + y. \end{cases}$
6. $\begin{cases} \frac{dx}{dt} = -x + 2y, \\ \frac{dy}{dt} = -2x - y. \end{cases}$
7. $\begin{cases} \frac{dx}{dt} = x + 2y, \\ \frac{dy}{dt} = -5x + y. \end{cases}$

By the Lyapunov function method examine for stability the stationary point $O(0, 0)$ of the systems:

8. $\begin{cases} \frac{dx}{dt} = -2x + y + xy^2, \\ \frac{dy}{dt} = -7x - 2y - 7x^2y. \end{cases}$
9. $\begin{cases} \frac{dx}{dt} = -y + x^2y^3, \\ \frac{dy}{dt} = x - x^3y^2. \end{cases}$
10. $\begin{cases} \frac{dx}{dt} = y + x^3, \\ \frac{dy}{dt} = x - y^3. \end{cases}$

Examine for stability in the first (linear) approximation the stationary point $O(0, 0)$ of the systems:

11. $$\begin{cases} \dfrac{dx}{dt} = -2x + \sin y, \\ \dfrac{dy}{dt} = 5(e^x - 1) - y. \end{cases}$$

12. $$\begin{cases} \dfrac{dx}{dt} = 2x - y\cos y, \\ \dfrac{dy}{dt} = 3x - 2y - xy^2. \end{cases}$$

Answers

1. Asymptotically stable. **2.** Unstable. **3.** Stable. **4.** Stable node. **5.** Saddle. **6.** Stable focus. **7.** Centre. **8.** Asymptocially stable, $v = 7x^2 + y^2$. **9.** Stable, $v = x^2 + y^2$. **10.** Unstable, $v = x^2 - y^2$. **11.** Asymptotically stable. **12.** Unstable.

Chapter 21

Special Topics of Differential Equations*

21.1 Asymptotic Behaviour of Solutions of Differential Equations as $x \to \infty$

Consider the differential equation

$$y'' + q(x)y = 0. \tag{21.1}$$

Let $q(x)$ as $x \to +\infty$ have a positive limit that, without any loss of generality, can be assumed to be equal to unity. Then

$$q(x) = 1 + \alpha(x),$$

where $\alpha(x)$ tends to zero as $x \to +\infty$, and equation (21.1) becomes

$$y'' + y + \alpha(x)y = 0. \tag{21.2}$$

When $x \to +\infty$ we obtain the "limiting" equation

$$y'' + y = 0, \tag{21.3}$$

All of its solutions are $y = A\cos x + B\sin x$, they are bounded for $x \in (-\infty, +\infty)$. Therefore, it is quite natural to expect that solutions of (21.2) are also bounded as $x \to +\infty$.

Theorem 21.1. *If $\alpha(x)$ is a continuously differentiable function, such that*

$$|\alpha(x)| < \frac{a}{x}, \qquad |\alpha'(x)| < \frac{a}{x^2}, \tag{21.4}$$

for all sufficiently large x, where a is a positive constant, then any solution of (21.2) is bounded as $x \to +\infty$.

◄ We multiply all the terms of (21.2) by y' and integrate the result with respect to x from some positive number x_0, which will be appropriately chosen later, to x:

$$(y')^2 \Big|_{x=x_0}^{x=x} + (y^2)\Big|_{x=x_0}^{x=x} + 2\int_{x_0}^{x} \alpha(x)yy'\,dx = 0.$$

---

* The material of the chapter is of interest and importance in applications and so some acquaintance with it is quite desirable.

Integrating by parts the last term on the left-hand side gives

$$y'^2(x) - y'^2(x_0) + y^2(x) - y^2(x_0) + (\alpha y^2)\Big|_{x=x_0}^{x=x} - \int_{x_0}^{x} \alpha'(x)y^2(x)dx = 0,$$

hence

$$\begin{aligned} y^2(x) &\leqslant y'^2(x) + y^2(x) \\ &\leqslant C(x_0) + |\alpha(x)|y^2(x) + \int_{x_0}^{x} |\alpha'(x)|y^2(x)dx, \end{aligned} \tag{21.5}$$

where $C(x_0) \geqslant 0$ is an expression dependent only on x_0.

We then denote by M the largest value of the function $|y(x)|$ on the interval $[x_0, x]$. Let it be achieved at a point $\xi \in [x_0, x]$. Using (21.4) and (21.5), we find

$$M^2 \leqslant C(x_0) + \frac{M^2 a}{\xi} + M^2 a\left(\frac{1}{x_0} - \frac{1}{\xi}\right); \quad M^2\left(1 - \frac{a}{x}\right) \leqslant C(x_0).$$

If we choose $x_0 \geqslant 2a$, we get

$$M^2 \leqslant 2C(x_0),$$

which proves the statement, since the quantity $2C(x_0)$ is independent of x. ⊳

Problem. Show that all the solutions of the equation

$$y'' + \left(1 + e^{-x^2} - \frac{1}{x+2}\right)y = 0$$

are bounded on $[0, +\infty)$.

If we impose on $\alpha(x)$ stronger conditions of decreasing:

$$\alpha(x) = O(1/x^2), \qquad x \to +\infty, \tag{21.6}$$

then, since (21.2) is more close to the limiting equation (21.3), solutions will be bounded, and also they will asymptomatically tend to trigonometric functions which are the solutions of the limiting equation. It can be shown that for any solution $y(x)$ of (21.2) holds the asymptotic formula

$$y(x) = A\sin(x + \delta_0) + O(1/x),$$

where A_1, δ_0 are some constants.

So the equation

$$y'' + \left(1 - \frac{\nu^2 - 1/4}{x^2}\right)y = 0,$$

where $\alpha(x) = \dfrac{1/4 - \nu^2}{x^2}$ satisfies (21.6). The solution to this equation is

related to the Bessel function $J_\nu(x)$ as $y(x) = \sqrt{x}J_\nu(x)$, which leads to an asymptotic formula for the Bessel function

$$J_\nu(x) = \frac{A}{\sqrt{x}}\sin(x+\delta_0) + O\left(\frac{1}{x^{3/2}}\right)$$

$$\left(\text{where } A = \sqrt{\frac{2}{\pi}}, \quad \delta_0 = -\frac{\pi\nu}{2} + \frac{\pi}{4}\right).$$

Examples show that the asymptotic behaviour of the differential equation's solution does not always follow from the behaviour of the limiting equations.

Consider, for example, the two equations

$$y'' - \frac{2}{x}y' + y = 0, \tag{21.7}$$

$$y'' + \frac{2}{x}y' + y = 0. \tag{21.8}$$

The limiting equation as $x \to +\infty$ for these equations is

$$y'' + y = 0. \tag{21.9}$$

All the solutions of (2.9) are bounded in $[1, +\infty]$. Equation (21.7) has the fundamental set of solutions

$$y_1(x) = \sin x - x\cos x, \qquad y_2(x) = \cos x + x\sin x,$$

whence it can be seen that all its nontrivial solutions are not bounded as $x \to +\infty$. On the other hand, (21.8) has the fundamental set of solutions

$$y_1(x) = \frac{\sin x}{x}, \qquad y_2(x) = \frac{\cos x}{x}.$$

All these solutions are bounded on $[1, +\infty)$ and even tend to zero as $x \to +\infty$.

21.2 Perturbation Method

A powerful method in applied mathematics is the method of the *small parameter* (or the *perturbation method*). Suppose that the formulation of a problem involves some small parameter ε together with the main unknows, and that the problem can be solved either exactly or nearly exactly at $\varepsilon = \varepsilon_0$. Without loss of generality, it can be assumed that $\varepsilon_0 = 0$. Then the solution to the problem for ε close to zero can in many cases be obtained as an expansion in powers of ε. The first term of the expansion will not contain ε and will be the unperturbed solution at $\varepsilon = 0$. Subsequent terms will be corrections to the perturbation of the solution.

We will illustrate the method by an example. We wish to solve the Cauchy problem

$$\frac{dy}{dx} = \frac{x}{1 + 0.1xy}, \qquad y(0) = 0. \tag{21.10}$$

The problem contains no parameters.

Consider a more general problem

$$\frac{dy}{dx} = \frac{x}{1 + \varepsilon xy}, \qquad y(0) = 0 \tag{21.11}$$

from which (21.10) follows at $\varepsilon = 0.1$. Problem (21.11) can easily be solved at $\varepsilon = 0$, its solution is $y = x^2/2$ (unperturbed solution).

We still look for a solution of problem (21.11) in the form of a series in powers of ε:

$$y(x, \varepsilon) = y_0(x) + \varepsilon y_1(x) + \varepsilon^2 y_2(x) + \dots, \tag{21.12}$$

hence

$$y'(x, \varepsilon) = y_0'(x) + \varepsilon y_1'(x) + \varepsilon^2 y_2'(x) + \dots .$$

We substitute $y(x, \varepsilon)$ and $y'(x, \varepsilon)$ into (21.11) and multiply by the denominator:

$$\begin{aligned}&[y_0'(x) + \varepsilon y_1'(x) + \varepsilon^2 y_2'(x) + \dots]\\ &\quad \times [1 + \varepsilon x y_0(x) + \varepsilon^2 x y_1(x) + \varepsilon^3 x y_2(x) + \dots] - x = 0. \end{aligned} \tag{21.13}$$

From the initial condition $y(0) = 0$ we have for (21.12)

$$y_0(0) + \varepsilon y_1(0) + \varepsilon^2 y_2(0) + \dots = 0.$$

Since ε is arbitrary, we obtain

$$y_0(0) = 0, \qquad y_1(0) = 0, \qquad y_2(0) = 0, \ \dots . \tag{21.14}$$

Because (21.13) must be obeyed for all sufficiently small $|\varepsilon|$, and the sequence of powers of ε is linearly independent, the coefficient at each power of ε vanishes independently. We remove the parentheses in (21.13) and equate to zero the coefficients at powers of ε:

$$\begin{array}{l|l} \varepsilon^0 & y_0' = x \text{ (unperturbed equation)}, \\ \varepsilon & y_1' + x y_0 y_0' = 0, \\ \varepsilon^2 & y_2' + x y_0 y_1' + x y_0' y_1 = 0, \\ & \text{etc.} \end{array} \tag{21.15}$$

From the first of (21.15) we, by (21.14), find $y_0 = x^2/2$. Substituting y_0 and y_0' into the second of (21.15), we obtain $y_1 = -x^5/10$. From the third equation we will then find $y_2 = 7x^8/160$, and so on. Equation (21.12) thus

becomes

$$y(x, \varepsilon) = \frac{x^2}{2} - \varepsilon \frac{x^5}{10} + \varepsilon^2 \frac{7x^8}{160} + \dots .$$

Specifically, for (21.10) we will get ($\varepsilon = 0.1$)

$$y = \frac{x^2}{2} - \frac{x^5}{100} + \frac{7x^8}{16\,000} + \dots$$

The series on the right converges quickly for $|x| < 1$.

Let us look at a more general situation. Consider the Cauchy problem for the equation

$$\frac{dy}{dx} = f(x, y, \varepsilon), \qquad y(0) = y^0, \tag{21.16}$$

where $\varepsilon > 0$ is a small parameter.

It can be shown that if (1) a function $f(x, y, \varepsilon)$ is continuous and has continuous and uniformly bounded derivatives of any order in the collection of arguments for $0 \leqslant x \leqslant a$, $-\infty < y < +\infty$, $0 \leqslant \varepsilon \leqslant \varepsilon_0$; (2) the Cauchy problem (21.16) at $\varepsilon = 0$

$$\frac{dy}{dx} = f(x, y, 0), \qquad y(0) = y^0, \tag{21.17}$$

has the only solution $y = \varphi(x)$, then in a certain interval $0 \leqslant x \leqslant x_0$ the Cauchy problem (21.16) has the solution $y = y(x, \varepsilon)$ which can be represented as

$$y(x, \varepsilon) = \sum_{k=0}^{\infty} \varepsilon^k y_k(x), \tag{21.18}$$

where $y_k(x)$ ($k = 0, 1, 2, \dots$) are unknown functions. We then for any $N \geqslant 1$ have the asymptotic formula

$$y(x, \varepsilon) = \sum_{k=0}^{N} \varepsilon^k y_k(x) + z_{N+1}(x, \varepsilon),$$

where $z_{N+1}(x, \varepsilon) \to 0$ as $\varepsilon \to 0$, $0 \leqslant x \leqslant x_0$, $z_{N+1}(x, \varepsilon) = O(\varepsilon^{N+1})$ for fixed N and $\varepsilon \to 0$.

To find $y_k(x)$, $k = 0, 1, 2, \dots$, we proceed as follows. We substitute the expression for y in the form of a formal series

$$y = y_0(x) + \varepsilon y_1(x) + \dots \tag{21.19}$$

into the equation (21.16):

$$\frac{dy_0}{dx} + \varepsilon \frac{dy_1}{dx} + \dots = f(x, y_0(x) + \varepsilon y_1(x) + \dots, \varepsilon). \tag{21.20}$$

Expanding the right-hand side of (21.20) also formally in powers of ε, we obtain

$$\frac{dy_0}{dx} + \varepsilon \frac{dy_1}{dx} + \ldots = f(x, y_0, 0) + \frac{\partial f(x, y_0, 0)}{\partial y} \varepsilon y_1 + \frac{\partial f(x, y_0, 0)}{\partial \varepsilon} \varepsilon + \ldots.$$

From the initial condition we, by (21.18), have

$$y_0(0) + \varepsilon y_1(0) + \ldots = y^0.$$

We then equate the terms at the same powers of ε

$$\frac{dy_0}{dx} = f(x, y_0, 0), \qquad y_0(0) = y^0,$$

$$\frac{dy_1}{dx} = \frac{\partial f(x, y_0, 0)}{\partial y} y_1 + \frac{\partial f(x, y_0, 0)}{\partial \varepsilon}, \qquad y_1(0) = 0,$$

....

We obtain for $y_0(x)$ the Cauchy problem (21.17), that, as assumed, has the only solution $y = \varphi(x)$, therefore $y_0 = \varphi(x)$. And for $y_1(x)$, we obtain the Cauchy problem

$$\frac{dy_1}{dx} = \frac{\partial f(x, \varphi(x), 0)}{\partial y} y_1(x) + \frac{\partial f(x, \varphi(x), 0)}{\partial \varepsilon}, \qquad y_1(0) = 0. \quad (21.21)$$

Equation (21.21) is a linear differential equation of the first order in $y_1(x)$, which is explicitly integrable, so that the function $y_1(x)$ will be constructed. For further approximations $y_2(x)$, $y_3(x)$, ... we will again obtain linear differential equations ($y_i(0) = 0$, $i = 2, 3, \ldots$), which are explicitly integrable as well.

The above procedure of constructing the solution $y(x, \varepsilon)$ provided good agreement with experiment, but for a long time it had not been theoretically justified. It was not before the 1930s that it was proved for a certain class of problems. Under certain conditions series (21.18) for $y(x, \varepsilon)$ converges, i.e., $y(x, \varepsilon)$ analytically depends on ε. In other cases this series appears to be convergent asymptotically.

The topic considered in outline above is known as *regular perturbations*, when the right-hand side $f(x, y, \varepsilon)$ of (21.16) is a sufficiently smooth function in y and ε. At present ever more important role is played by the theory of *singular perturbations*.

21.3 Oscillations of Solutions of Differential Equations

Both from fundamental and practical points of view it is very important to be able to solve the question of the presence of zeros of the solution $y(x)$ of the equation

$$p_0(x)y'' + p_1(x)y' + p_2(x)y = 0 \tag{21.22}$$

in the interval (a, b), i.e., those values of $x \in (a, b)$ at which the solution $y(x)$ becomes zero. Consider the simplest equation of the second order with constant coefficients

$$y'' + qy = 0, \qquad q = \text{const}.$$

If $q \leqslant 0$, then each solution of this equation can in the entire interval $-\infty < x < +\infty$ become zero at no more than one point. For $q > 0$ each solution

$$y = C_1 \cos \sqrt{q}\, x + C_2 \sin \sqrt{q}\, x = A \sin(\sqrt{q}\, x + \delta)$$

has an infinite number of zeros separated by $\pi/\sqrt{q}$, i.e., the spacing is the smaller the larger q.

Definition. A solution $y(x)$ of a differential equation is said to be *nonoscillating* in a given interval, if it has in the interval no more than one zero. Otherwise, the solution is called *oscillating*.

An equation of the form $y'' + qy = 0$ (q = const) has thus nonoscillating solutions in any interval, if $q \leqslant 0$, and oscillating solutions in a sufficiently large interval, if $q > 0$.

We can now generalize this result to equations of the second order with variable coefficients. Suppose that the coefficients of the equation are real and examine *only real solutions* of such equations. Consider the equation

$$y'' + q(x)y = 0 \tag{21.23}$$

to which any equation (21.22) comes down.

Theorem 21.2. *If* $q(x) \leqslant 0$ *everywhere in the interval* (a, b), *then all the solutions of* (21.23) *are nonoscillating in the interval* (a, b).

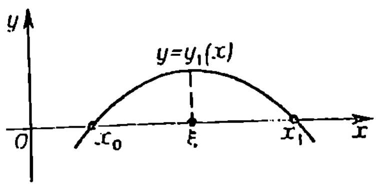


Fig. 21.1

The theorem has a geometrical explanation. Suppose that some solution $y_1(x) \not\equiv 0$ of (21.23) has at least two zeros in the interval (a, b). Let them

be x_0 and x_1, $x_0 < x_1$, and let $y_1(x)$ have no other zeros in the interval (x_0, x_1) (Fig. 21.1). Being a continuous function, $y_1(x)$ will then retain a constant sign in the interval (x_0, x_1). To be more specific, we will suppose that $y_1(x) > 0$ in (x_0, x_1) (otherwise we would take the solution $-y_1(x)$).

At a point $\xi \in (x_0, x_1)$ the function $y_1(x)$ will have a positive maximum. Correspondingly, in a certain neighbourhood of the point ξ we will have $y''(x) < 0$. On the other hand, if $q(x) \leqslant 0$ in (a, b), then from equation (21.23) it follows that $y_1'' \geqslant 0$ everywhere in (x_0, x_1). The resultant contradiction suggests that our assumption was wrong and all the solutions of the equation are nonoscillating.

Theorem 21.3 (Sturm's separation theorem). *If x_0 and x_1 are two successive zeros of the solution $y_1(x)$ of the differential equation* (21.23), *then any other linearly independent solution $y_2(x)$ of the equation has just one zero between x_0 and x_1*. In other words, zeros of two linearly independent solutions of (21.23) mutually separate each other.

◄ Suppose that in the entire interval (x_0, x_1) the solution $y_2(x)$ has no zeros. The solutions $y_1(x)$ and $y_2(x)$ being linearly independent, $y_2(x)$ does not vanish at x_0 and x_1. We form the Wronskian

$$W(x) = y_1'(x)y_2(x) - y_2'(x)y_1(x). \tag{21.24}$$

Since $W(x)$ is not zero, it retains a constant sign. Suppose for definiteness that $W(x) > 0$. Dividing both sides of (21.24) by $y_2^2(x)$ gives

$$\frac{y_1'(x)y_2(x) - y_2'(x)y_1(x)}{y_2^2(x)} = \frac{W(x)}{y_2^2(x)},$$

or

$$\frac{d}{dx}\left(\frac{y_1(x)}{y_2(x)}\right) = \frac{W(x)}{y_2^2(x)}.$$

By the assumption that $y_2(x) \neq 0$ on $[x_0, x_1]$ we have on the right-hand side a continuous function of x. Integrating the last identity with respect to x from x_0 to x_1, we will have

$$\left.\frac{y_1(x)}{y_2(x)}\right|_{x=x_0}^{x=x_1} = \int_{x_0}^{x_1} \frac{W(x)}{y_2^2(x)}\,dx.$$

The left-hand side is zero by the condition that $y_1(x_0) = y_1(x_1) = 0$, and the right-hand side contains an integral of a positive function, i.e., it is a positive quantity. The contradiction proves that between two successive zeros of $y_1(x)$ there exists at least one zero $y_2(x)$. If there were two zeros $y_2(x_3) = y_2(x_4) = 0$, $x_0 < x_3 < x_4 < x_1$, then, interchanging $y_1(x)$ and $y_2(x)$, we would prove that there exists a zero of the function $y_1(x)$ between

x_3 and x_4, and hence between x_0 and x_1. But this is at variance with the condition that $y_1(x)$ has no zeros between x_0 and x_1. ►

For example, for the equation $y'' + y = 0$ the functions $y_1 = \sin x$, $y_2 = \cos x$ are two linearly independent solutions. Their zeros mutually separate each other. The equation also has the complex-valued solution $y = \cos x + i \sin x$, which has no zeros in any interval of the real axis.

Problem. Show that if in the interval (a, b) one solution of the equation $y'' + q(x)y = 0$ has more than two zeros, then all the solutions are oscillating.

Theorem 21.4 (comparison theorem). *Suppose we have two equations*

$$y'' + q_1(x)y = 0 \tag{21.25}$$

and

$$z'' + q_2(x)z = 0. \tag{21.26}$$

If $q_2(x) \geqslant q_1(x)$ in the interval (a, b), then between each two zeros of any solution $y(x)$ of (21.25) *there is at least one zero of each solution $z(x)$ of* (21.26).

The comparison theorem is normally applied when either (21.25) or (21.26) is an equation with constant coefficients. Consider the equation

$$y'' + q(x)y = 0, \tag{21.27}$$

where $q(x) > 0$ on the interval $[a, b]$ and the function $q(x)$ is continuous on it, and let $M = \max_{a \leqslant x \leqslant b} q(x)$, $m = \min_{a \leqslant x \leqslant b} q(x)$. Suppose now that $M > m$, so that $q(x) \not\equiv \text{const}$ on $[a, b]$. If then we take (21.25) to be $y'' + my = 0$ and (21.26) to be (21.27), we will find that the distance between two successive zeros of the solution of (21.27) is smaller than $\pi/\sqrt{m}$. If then we take the original equation (21.27) to be (21.25), and take (21.26) to be $y'' + My = 0$, we will conclude that the distance between two successive zeros of the solution of (21.27) is not smaller than $\pi/\sqrt{M}$.

This theorem estimates from above and below the separation between zeros of oscillating solutions of differential equations. It can also be shown that if $\lim_{x \to \infty} q(x) = q > 0$, then any solution of (21.27) is infinitely oscillating, the separation betwen zeros tending to $\pi/\sqrt{q}$.

For example, for the Bessel equation

$$x^2y'' + xy' + (x^2 - \nu^2)\,y = 0, \qquad x > 0,$$

putting $y = x^{-1/2}z$ gives

$$z'' + \left(1 - \frac{\nu^2 - 1/4}{x^2}\right)z = 0.$$

At sufficiently large x the expression $1 - (\nu^2 - 1/4)/x^2$ can be made arbitrarily close to unity. Therefore, for sufficiently large values of x the spacing between consequtive zeros of the solutions of the Bessel equation is arbitrarily close to π.

Problem. Show that as x increases the successive zeros of any solution of Airy's equation

$$y'' + xy = 0, \qquad x > 0,$$

come indefinitely close. (Airy's equation occurs in various applications, e.g., in quantum mechanics, and it is not integrable using elementary methods.)

Exercises

1. Show that in $[0, +\infty)$ all the solutions of the equation

$$y'' + \left(1 + \frac{1}{1 + x^2}\right) y = 0$$

are bounded.

2. Derive the asymptotic formula with the remainder $O(\varepsilon^2)$ for the solution $y(x, \varepsilon)$ of the Cauchy problem for the Riccati equation

$$\frac{dy}{dx} = 2xy + e^{-x^2} + \varepsilon xy^2, \qquad y(0) = 0.$$

3. Consider the Cauchy problem $y' + y = \varepsilon y^2$, $y(0) = 1$:

(a) find three expansion terms for the solution when the small quantity $\varepsilon > 0$;

(b) show that the exact solution has the form

$$y = \frac{e^{-x}}{1 + \varepsilon(e^{-x} - 1)};$$

(c) expand this exact solution for small $\varepsilon > 0$ and compare the results with those of (a).

4. Find three or four expansion terms for the solution in powers of small parameter ε:

(a) $y' = 4\varepsilon x + y^2$, $y(1) = 1$; (b) $y' = e^{y-x} + \varepsilon y$, $y(0) = -\varepsilon$.

Answers

2. $y = xe^{x^2} + \frac{\varepsilon}{2} e^{x^2}[1 + e^{x^2}(x^2 - 1)] + O(\varepsilon^2)$. 4. (a) $y = \frac{1}{x} + \varepsilon\left(x^2 - \frac{1}{x^2}\right) + O(\varepsilon^2)$; (b) $y = x - \varepsilon(x + 1) + \frac{\varepsilon^2}{2}\left(e^x - x^2 - 2x - 1\right) + O(\varepsilon^3)$.

Chapter 22

Multiple Integrals. Double Integral

22.1 Problem Leading to the Concept of Double Integral

We come to the notion of double integral when handling the specific geometry problem of computing the volume of a cylindrical body.

A *cylindrical body* is a body bounded by the xy-plane, some surface $z = f(x, y)$ and a cylindrical surface whose generatrices are parallel to the axis (Fig. 22.1). A domain D is called the *base* of the cylindrical body. This is the orthogonal projection on the xy-plane of the surface that bounds the cylindrical body.

To determine the volume we will follow the two principles:

(1) if we break a body into parts, its volume will be equal to the sum of the volumes of the constituent parts (*additivity*).

(2) the volume of a right cylinder bounded by the plane $z = \text{const}$ parallel to the xy-plane is equal to the area of the base multiplied by the height of the body.

In what follows we will assume that the domain D is squarable (i.e., having an area) and bounded (i.e., lying within a circle whose centre is the origin of coordinates).

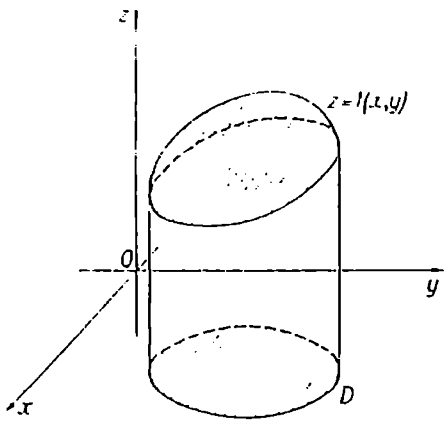


Fig. 22.1

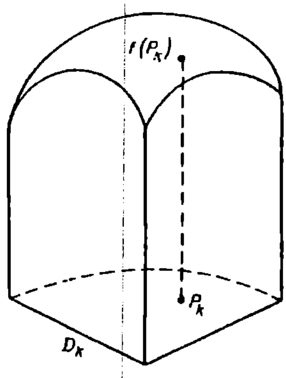


Fig. 22.2

Let $z = f(x, y)$ be the equation of a surface that bounds the cylindrical body, and let $f(x, y)$ be a continuous function of a point $P(x, y)$ in D. To begin with, we assume that the surface wholly lies above the xy-plane, i.e., that $f(x, y) \geqslant 0$ everywhere in D.

We denote the volume of the cylindrical body by V.

We now break the base D of the cylindrical body into n nonintersecting squarable domains of arbitrary shape (we call them *partial domains* or *regions*).

We will assign some numbers to the partial regions and denote them by $D_1, D_2, \ldots, D_n$, and their respective areas by $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$. We will introduce the notion of the *diameter* of D_k:

$$\operatorname{diam} D_k = \sup_{P, Q \in D_k} \varrho(P, Q),$$

where $\varrho(P, Q)$ is the distance between the points P and Q. We further denote by d the largest of the diameters of D_k ($k = 1, 2, \ldots, n$). Through the boundary of each partial region we will draw a cylindrical surface with generatrices parallel to the z-axis. As a result, the cylindrical body will be broken into n partial cylindrical bodies.

We replace the kth partial body by a right cylinder with the same base and height equal to the z-coordinate of some point on the surface to be replaced (Fig. 22.2).

The volume of such a cylinder will be

$$\Delta V_k = f(P_k)\Delta S_k,$$

where the point $P_k(x_k, y_k) \in D_k$, and ΔS_k is the area of D_k.

After we have subjected each partial cylindrical body to this procedure, we end up with an n-stepped body of volume

$$V_n = \sum_{k=1}^{n} f(P_k)\Delta S_k. \tag{22.1}$$

It is intuitively clear that V_n represents the volume V we seek for the more accurately the smaller are the sizes of D_k.

By definition, we assume V to be the limit to which the volume of the n-stepped body (22.1) tends when $n \to \infty$ and when the largest of the diameters d or D_k tends to zero. It is only natural that the limit should not depend on the way in which D has been broken up into partial domains D_k and points P_k in D_k have been selected.

Let us take the general case. The sum (22.1) is called the *integral sum* for $f(x, y)$ in D for a given division of D into n partial domains and a given selection of the points $P_k(x_k, y_k)$ within D_k. We make n tend to infinity so that $d \to 0$. The sum (22.1) will then change.

Definition. If there exists a limit of integral sums $\sigma = \sum_{k=1}^{n} f(P_k)\Delta S_k$ as $d \to 0$, such that it depends neither on the choice of division of D into partial domains, nor on the selection of points P_k within D_k, then it is called the *double integral* of $f(P)$, or $f(x, y)$, over D, and is denoted by $\iint\limits_D f(P)ds$ or $\iint\limits_D f(x, y)dx\,dy$. Thus,

$$\iint\limits_D f(P)ds = \lim_{d\to 0} \sum_{k=1}^{n} f(P_k)\Delta S_k. \tag{22.2}$$

Function $f(x, y)$ is then said to be *integrable over the domain D*. Henceforth d is the largest of the diameters of partial domain D_k, $f(P)$ is the integrand, $f(P)ds$ is the integration element, ds is the differential (or element) of area, D is the domain of integration, $P(x, y)$ is the variable point of integration.

Returning to the definition of the volume of a cylindrical body, we conclude that the volume of the cylindrical body bounded by the xy-plane, the surface $z = f(x; y)$ $(f(x, y) \geqslant 0)$ and the cylindrical surface whose generatrix is parallel to the z-axis is determined by the double integral of $f(x, y)$ over the domain D, which is the base of the body:

$$V = \iint\limits_D f(P)ds \qquad \text{or} \qquad V = \iint\limits_D f(x, y)dx\,dy.$$

Here $dx\,dy$ is a surface element in Cartesian coordinates.

So much for the geometrical meaning of the double integral of a non-negative function.

If $f(P) \leqslant 0$ in D, then $V = -\iint\limits_D f(P)ds$.

If in D the function $f(P)$ assumes both positive and negative values, then the integral $\iint\limits_D f(P)ds$ is the algebraic sum of the volumes of those parts of the body that lie above the xy-plane (they are taken with the plus sign) and those parts that lie under the xy-plane (they are taken with the minus sign).

Sums of the form (22.1) are set up for a function of two independent variables and a limiting transition is then carried out in a wide variety of problems, for example, when calculating the mass of a plane figure.

It is worth noting that the condition for $f(x, y)$ to be bounded in D is not sufficient for the function to be integrable. For example, take the function $f(x, y)$ defined in a square D: $\{0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$ in the following manner: $f(x, y) = 1$, if x and y are rational, and $f(x, y) = 0$ otherwise. This function is bounded but not integrable in the above sense.

We now turn to the sufficient condition for integrability.

Theorem 22.1. *Any function $f(x, y)$ that is continuous in a bounded closed domain D is integrable over that domain.*

The requirement that the integrand be continuous is too constringent. Therefore, for applications we have the following important theorem that guarantees the existence of the double integral for a certian class of discontinuous functions.

Theorem 22.2. *If a function $f(x, y)$ is bounded in a bounded closed domain D and continuous everywhere in D, save for a set of points of zero surface area*[*)]*, then it is integrable over D.*

22.2 Main Properties of Double Integral

Double integrals have a number of properties similar to those of the definite integrals for a function of one independent variable.

(1) *Linearity.* If the functions $f(P)$ and $\varphi(P)$ are integrable over a domain D, and α and β are any real numbers, then the function $\alpha f(P) + \beta\varphi(P)$ is also integrable over D, and

$$\iint_D [\alpha f(P) + \beta\varphi(P)]ds = \alpha \iint_D f(P)ds + \beta \iint_D \varphi(P)ds. \tag{22.3}$$

(2) If the functions $f(P)$ and $\varphi(P)$ are integrable over a domain D and everywhere in D $f(P) \leqslant \varphi(P)$, then

$$\iint_D f(P)ds \leqslant \iint_D \varphi(P)ds, \tag{22.4}$$

i.e., inequalities can be integrated *term by term*. Specifically, integrating the inequalities

$$-|f(P)| \leqslant f(P) \leqslant |f(P)|,$$

we will get

$$-\iint_D |f(P)|ds \leqslant \iint_D f(P)ds \leqslant \iint_D |f(P)|ds,$$

or similarly,

$$\left|\iint_D f(P)\,ds\right| \leqslant \iint_D |f(P)|ds.$$

(3) *Area of a plane region.* The area of a plane domain D is equal to the double integral over that domain of the constant function identically equal to unity.

[*)] We will say that a set of points in a plane, in particular a curve, has a *zero area*, if it can be contained by a polygonal figure of an infinitely small area.

Indeed, the integral sum of $f(P) \equiv 1$ in D has the form

$$\sum_{k=1}^{n} 1 \cdot \Delta S_k$$

and for any division of D into D_k it is equal to the area S. But then the limit of that sum, i.e., the double integral, will be equal to the area S of the domain D:

$$S = \iint\limits_D ds. \tag{22.5}$$

(4) *Estimation of integral.* Suppose that a function $f(P)$ is continuous in a bounded closed domain D. Let M and m be respectively the largest and smallest values of $f(P)$ in D, and S be the area of D. Then

$$mS \leqslant \iint\limits_D f(P)ds \leqslant MS. \tag{22.6}$$

(5) *Additivity.* If a function $f(P)$ is integrable over a domain D and the domain D is divided into two domains D_1 and D_2 such that they have no common internal points, then $f(P)$ is integrable over each of D_1 and D_2, so that

$$\iint\limits_D f(P)ds = \iint\limits_{D_1} f(P)ds + \iint\limits_{D_2} f(P)ds. \tag{22.7}$$

(6) **Theorem 22.3 (mean value theorem).** *If a function $f(P)$ is continuous in a closed bounded domain D, then there exists at least one point P_m in D such that*

$$\iint\limits_D f(P)ds = f(P_m)S, \tag{22.8}$$

where S is the area of D.

◄ Since $f(P)$ is continuous in the closed bounded domain D, it assumes in D a maximum value M and a minimum value m. By property (4) we have

$$mS \leqslant \iint\limits_D f(P)ds \leqslant MS,$$

hence

$$m \leqslant \frac{1}{S}\iint\limits_D f(P)ds \leqslant M.$$

The number $\frac{1}{S}\iint\limits_D f(P)ds$ is thus contained between the maximum and minimum values of $f(P)$ in D. The function $f(P)$ being continuous in D

assumes at some point*) $P_m \in D$ a value equal to the number $\frac{1}{S}\iint\limits_D f(P)ds$:

$$f(P_m) = \frac{1}{S}\iint\limits_D f(P)ds. \tag{22.9}$$

Whence we have

$$\iint\limits_D f(P)ds = f(P_m)S. \quad ►$$

The value $f(P_m)$ defined by (22.9) is known as the *mean value of $f(P)$ in D*.

The mean value theorem has a geometrical meaning.

If in a domain D a function $f(P) \geqslant 0$ is defined, then (22.8) implies that there exists a right cylinder with a base D (of area S) and height $H = f(P_m)$ whose volume is equal to the volume of the cylindrical body (Fig. 22.3).

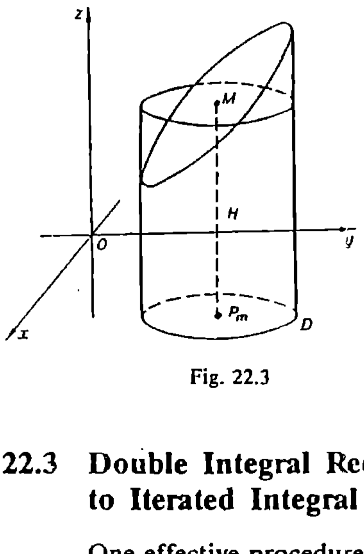


Fig. 22.3

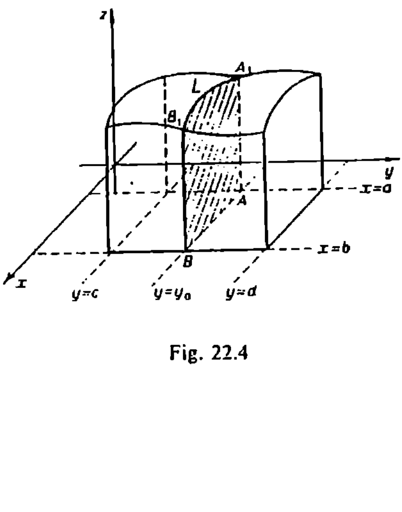


Fig. 22.4

22.3 Double Integral Reduced to Iterated Integral

One effective procedure to take a double integral is to reduce it to an iterated one.

The case of a rectangle. Let a domain D be a closed rectangle Π with sides parallel to the coordinate axes

$$\Pi = \{a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d\}.$$

---

*) By definition a domain is a connected set.

Suppose now that a function $f(x, y)$ is continuous in Π. The double integral

$$\iint_{\Pi} f(x, y)\,dx\,dy$$

can be treated as the (algebraic) volume of the cylindrical body with the base Π, bounded by the surface $z = f(x, y)$.

Consider the cylindrical body. We will draw the plane $y = y_0$, $c \leqslant y_0 \leqslant d$, perpendicular to the y-axis (Fig. 22.4). This plane will cut the body through so that the cross-section will be the curvilinear trapezoid ABB_1A_1 bounded above by the plane line L: $z = f(x, y_0)$, $y = y_0$.

The area of the trapezoid is given by the integral

$$\int_a^b f(x, y_0)\,dx, \tag{22.10}$$

where the integration is carried out with respect to x and y_0 is the second argument of the integrand, which in this case is treated as a constant ($c \leqslant y_0 \leqslant d$). The value of (22.10) depends on the choice of the value of y_0. We write

$$S(y) = \int_a^b f(x, y)\,dx. \tag{22.11}$$

Expression (22.11) gives the area of the cross-section of the cylindrical body as a function of y. Therefore, the volume of the body can be computed by

$$V = \int_c^d S(y)\,dy,$$

On the other hand, this volume is expressed by the double integral of $f(x, y)$ over the rectangle Π. Hence

$$\iint_{\Pi} f(x, y)\,dx\,dy = \int_c^d S(y)\,dy.$$

Substituting (22.11) gives

$$\iint_{\Pi} f(x, y)\,dx\,dy = \int_c^d \left(\int_a^b f(x, y)\,dx\right) dy.$$

This relation is generally written as

$$\iint_{\Pi} f(x, y)\,dx\,dy = \int_c^d dy \int_a^b f(x, y)\,dx. \tag{22.12}$$

The volume of a cylindrical body can also be worked out from the areas of flat cross-sections by planes $x = x_0$. This yields

$$\iint\limits_{\Pi} f(x, y)dx\,dy = \int_a^b dx \int_c^d f(x, y)dy. \tag{22.13}$$

Each of the expressions on the right sides of (22.12) and (22.13) contains two successive operations of conventional integration of $f(x, y)$. The relations are known as *iterated* (or *repeated*) *integrals of* $f(x, y)$ *over the domain* Π.

Comparing (22.12) and (22.13), we notice that if $f(x, y)$ is continuous in the closed rectangle Π, then it is always possible to pass to iterated integrals and

$$\int_a^b dx \int_c^d f(x, y)dy = \int_c^d dy \int_a^b f(x, y)dx, \tag{22.14}$$

i.e., values of iterated integrals of the function $f(x, y)$, which is continuous in Π, are independent of the order of integration.

Example. Find the double integral of $z = x^2 + y^2$ over the domain $\Pi = \{0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$.

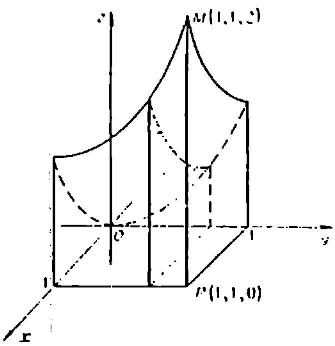


Fig. 22.5

◄ We have (see Fig. 22.5):

$$\iint\limits_{\Pi} (x^2 + y^2)dx\,dy = \int_0^1 dx \int_0^1 (x^2 + y^2)dy$$

$$= \int_0^1 \left(x^2 y + \frac{y^3}{3}\right)\Big|_{y=0}^{y=1} dx = \int_0^1 \left(x^2 + \frac{1}{3}\right)dx$$

$$= \left(\frac{x^3}{3} + \frac{x}{3}\right)\Big|_{x=0}^{x=1} = \frac{2}{3}. \quad ►$$

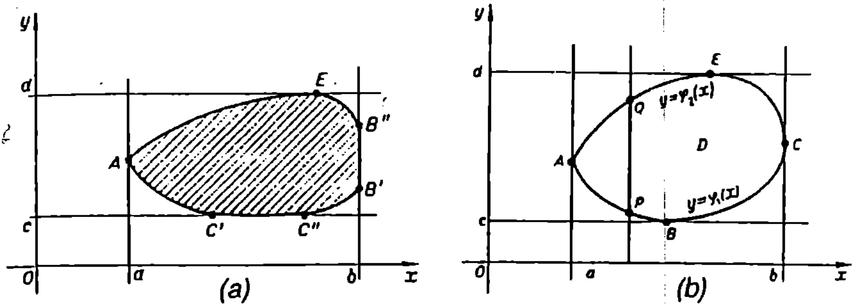


Fig. 22.6

The case of an arbitrary region. Suppose now that the domain of integration is an arbitrary bounded squarable closed domain D in the xy-plane. We further assume that D meets the following condition: any straight line x = const ($a \leqslant x \leqslant b$) parallel to the y-axis meets the boundary of D at no more than two points or over a segment (Fig. 22.6a).

We put some domain D inside the rectangle $\Pi = \{a \leqslant x \leqslant b, c \leqslant y \leqslant d\}$. The segment $[a, b]$ is an orthogonal projection of D on the x-axis, and the segment $[c, d]$ is an orthogonal projection of D on the y-axis. Points A and C divide the boundary of D into the curves ABC and AEC (Fig. 22.6b). Each of the curves meets an arbitrary straight line parallel to the y-axis at no more than one point. Therefore, their equations can be written in a form solvable for y:

$$\begin{aligned} &(ABC): \quad y = \varphi_1(x), \\ & \qquad\qquad\qquad\qquad\qquad (a \leqslant x \leqslant b). \\ &(AEC): \quad y = \varphi_2(x), \end{aligned} \tag{22.15}$$

We now cut the body with the plane x = const ($a < x < b$). The resultant cross-section will be the curvilinear trapezoid $PQMN$ (Fig. 22.7),

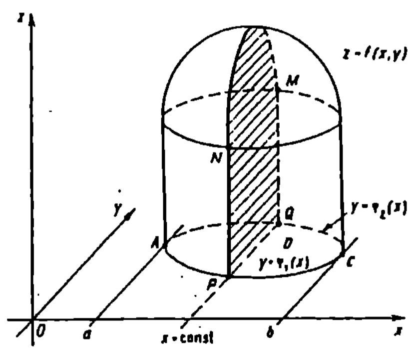


Fig. 22.7

whose area is given by a conventional integral of $f(x, y)$ viewed as a function of one variable y. Here y varies from the ordinate $\varphi_1(x)$ of the point P to the ordinate $\varphi_2(x)$ of the point Q. Point P is the point of "entrance" of the straight line x = const (in the xy-plane) into D, and Q is the point of "exit" from D. Since the equation of the curve ABC is $y = \varphi_1(x)$, and of the curve AEC it is $y = \varphi_2(x)$, then these ordinates will for a given x be $\varphi_1(x)$ and $\varphi_2(x)$, respectively. Consequently, the integral

$$\int_{\varphi_1 x}^{\varphi_2 x} f(x, y)dy = S(x) \tag{22.16}$$

gives us an expression for the area of a plane cross-section of a cylindrical body as a function of the position of the secant plane x = const. The volume of the body will be equal to the integral of that expression with respect to x in the range of x $(a \leqslant x \leqslant b)$.

Thus

$$\iint_D f(x, y)dx\,dy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)dy. \tag{22.17}$$

In particular, the surface area S of D will be

$$S = \iint_D dx\,dy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} dy. \tag{22.18}$$

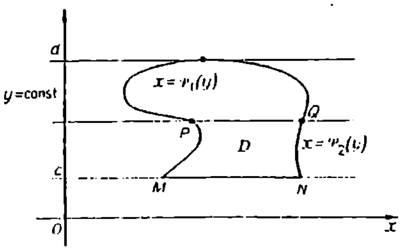


Fig. 22.8

We now suppose that each straight line y = const $(c \leqslant y \leqslant d)$ meets the boundary of the domain D at no more than two points P and Q, whose abscissas are $\psi_1(y)$ and $\psi_2(y)$, respectively; or over a segment MN (Fig. 22.8).

Similar arguments lead to the formula

$$\iint_D f(x, y)dx\,dy = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x, y)dx, \tag{22.19}$$

which also reduces the computation of a double integral to an iterated integral.

Example. Take the double integral of $f(x, y) = 2x - y + 3$ over a domain D bounded by the lines $y = x$ and $y = x^2$ (Fig. 22.9).

First method. Examine D. The straight line $y = x$ and the parabola $y = x^2$ meet at points $O(0, 0)$ and $M(1, 1)$. And so, x varies from 0 to

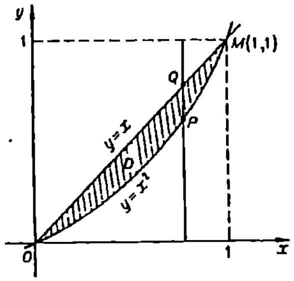


Fig. 22.9

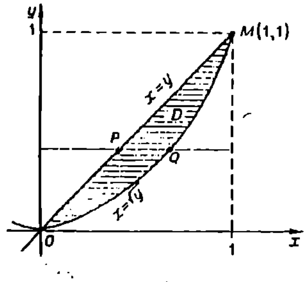


Fig. 22.10

1, and $\varphi_1(x) = x^2$ and $\varphi_2(x) = x$. Any straight line $x = \text{const}$ $(0 \leqslant x \leqslant 1)$ meets the boundary of the region at no more than two points. Therefore, we can here apply (22.17)

$$\iint\limits_D (2x - y + 3)\,dx\,dy = \int_0^1 dx \int_{x^2}^{x} (2x - y + 3)\,dy$$

$$= \int_0^1 \left(2xy - \frac{y^2}{2} + 3y\right)\Bigg|_{y=x^2}^{y=x} dx$$

$$= \int_0^1 \left(2x^2 - \frac{x^2}{2} + 3x - 2x^3 + \frac{x^4}{2} - 3x^2\right) dx$$

$$= \int_0^1 \left(3x - \frac{3}{2}x^2 - 2x^3 + \frac{x^4}{2}\right) dx = \frac{3}{5}.$$

Second method (Fig. 22.10). Using (22.19) gives

$$\iint\limits_D (2x - y + 3)\,dx\,dy = \int_0^1 dy \int_y^{\sqrt{y}} (2x - y + 3)\,dx$$

$$= \int_0^1 (x^2 - xy + 3x)\Bigg|_{x=y}^{x=\sqrt{y}} dy = \int_0^1 (3\sqrt{y} - 2y + y^{3/2})\,dy = \frac{3}{5}. \blacktriangleright$$

Example. Calculate the volume of a body bounded by the surface $z = 1 - 4x^2 - y^2$ and the xy-plane.

◀ The elliptic paraboloid $z = 1 - 4x^2 - y^2$ is intersected by the xy-plane along the line

$$L: \quad \begin{cases} z = 0 \\ z = 1 - 4x^2 - y^2, \end{cases}$$

$\cos^4 t = \frac{1}{8}(3+$

i.e.,

$$\begin{cases} z = 0, \\ 4x^2 + y^2 = 1, \end{cases}$$

or

$$\frac{x^2}{(1/2)^2} + \frac{y^2}{1} = 1, \qquad z = 0.$$

This is an ellipse with semi-axes $a = 1/2$ and $b = 1$ (Fig. 22.11).

The body being symmetrical about the coordinate axes, we get

$$V = 4\iint_D (1 - 4x^2 - y^2)\,dx\,dy = 4\int_0^{1/2} dx \int_0^{\sqrt{1-4x^2}} (1 - 4x^2 - y^2)\,dy$$

$$= 4\int_0^{1/2} \left[(1 - 4x^2)y - \frac{y^3}{3}\right]\Bigg|_{y=0}^{y=\sqrt{1-4x^2}} dx = \frac{8}{3}\int_0^{1/2} (1 - 4x^2)^{3/2}\,dx$$

$$= \left| \begin{array}{ll} 2x = \sin t, & 2dx = \cos t\,dt \\ 1 - 4x^2 = 1 - \sin^2 t = \cos^2 t & \\ x_1 = 0, & t_1 = 0 \\ x_2 = 1/2, & t_2 = \pi/2 \end{array} \right| = \frac{4}{3}\int_0^{\pi/2} (\cos^2 t)^{3/2} \cos t\,dt$$

$$= \frac{4}{3}\int_0^{\pi/2} \cos^4 t\,dt = \frac{4}{3}\int_0^{\pi/2} \frac{1 + 2\cos 2t + \cos^2 2t}{4}\,dt$$

$$= \frac{1}{3}(t + \sin 2t)\Big|_0^{\frac{\pi}{2}} + \frac{1}{3}\int_0^{\pi/2} \frac{1 + \cos 4t}{2}\,dt$$

$$= \frac{\pi}{6} + \frac{1}{6}\left(t + \frac{\sin 4t}{4}\right)\Bigg|_0^{\frac{\pi}{2}} = \frac{\pi}{6} + \frac{\pi}{12} = \frac{\pi}{4}. \blacktriangleright$$

Remark. If a domain D is such that some straight lines (vertical or horizontal) meet its boundary at more than two points, then to determine a double integral over D we should divide it appropriately into parts, reduce

each integral over these parts to an iterated integral and add up the results obtained.

Example. Take the double integral $\iint\limits_D e^{x+y}\,dx\,dy$ over a domain D contained between two squares with a common centre at the origin of coordinates and sides parallel to the coordinate axes, if for the internal square the side is 2, and for the external 4.

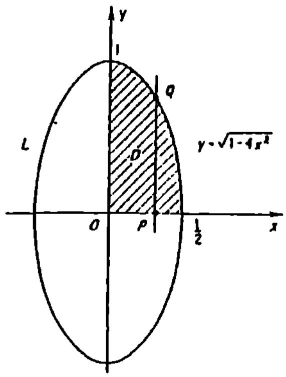


Fig. 22.11

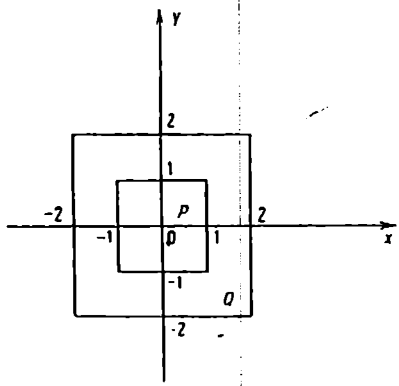


Fig. 22.12

◄ The function $f(x, y) = e^{x+y}$ is continuous both in the large (Q) and the small (P) squares (Fig. 22.12). According to Theorem 22.1 there exist integrals of e^{x+y} over the squares, and so the integral will be

$$\iint\limits_D e^{x+y}\,dx\,dy = \iint\limits_Q e^{x+y}\,dx\,dy - \iint\limits_P e^{x+y}\,dx\,dy,$$

where

$$\iint\limits_Q e^{x+y}\,dx\,dy = \int\limits_{-2}^{2}\left(\int\limits_{-2}^{2} e^{x+y}\,dy\right)dx = \int\limits_{-2}^{2} e^{x}\,dy \int\limits_{-2}^{2} e^{y}\,dy$$

$$= e^{x}\Big|_{-2}^{2} \cdot e^{y}\Big|_{-2}^{2} = (e^{2} - e^{-2})^{2},$$

$$\iint\limits_P e^{x+y}\,dx\,dy = \int\limits_{-1}^{1} e^{x}\,dx \int\limits_{-1}^{1} e^{y}\,dy = (e - e^{-1})^{2}.$$

Hence

$$\iint\limits_D e^{x+y}\,dx\,dy = (e^{2} - e^{-2})^{2} - (e - e^{-1})^{2}$$

$$= e^{4} - 2 + e^{-4} - e^{2} + 2 - e^{-2}$$

$$= 2\cosh 4 - 2\cosh 2.$$ ►

22.4 Change of Variables in Double Integral

Curvilinear coordinates. Suppose that in a domain D^* in the uv-plane we are given a pair of functions

$$\begin{cases} x = \varphi(u, v), \\ y = \psi(u, v), \end{cases} \tag{22.20}$$

which we will consider continuous and having continuous partial derivatives.

By (22.20), corresponding to each point $M^*(u, v)$ in D^* is one definite point $M(x, y)$ in the xy-plane, and so corresponding to the points of D^* is some set D of points (x, y) in the xy-plane (see Fig. 22.13a, b).

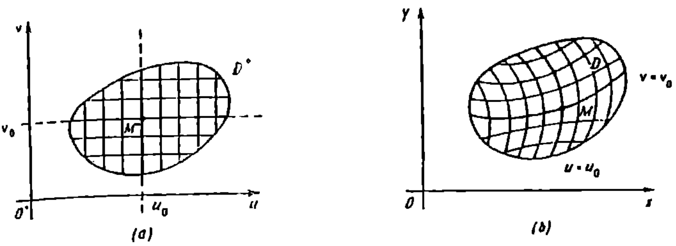


Fig. 22.13

We say then that the functions (22.20) map D^* into a set D. Suppose that corresponding to various points (u, v) are various points (x, y). This is equivalent to saying that equations (22.20) are uniquely solvable for u, v:

$$\begin{cases} u = g(x, y), \\ v = h(x, y). \end{cases} \tag{22.21}$$

In this case, the mapping is said to be a *one-to-one mapping* of domains D and D^*. Under such a mapping any continuous curve L^* in D^* will be mapped into a continuous curve L in D.

If the functions $g(x, y)$ and $h(x, y)$ are continuous, then any continuous curve $L \subset D$ will by (22.11) be mapped into the continuous line $L^* \subset D^*$.

Since from a given pair of values (u_0, v_0) of variables u, v in D^* we can uniquely determine not only the position of the point $M^*(u_0, v_0)$ in the domain D^* itself, but also the position of the corresponding point $M(x_0, y_0)$ in D ($x_0 = \varphi(u_0, v_0)$, $y_0 = \psi(u_0, v_0)$), this will enable us to consider the numbers u and v as some new coordinates of M in the xy-plane. They are called the *curvilinear coordinates of the point M.*

The set of points in D such that one of their coordinates maintains a constant value is called the *coordinate line*. Putting in (22.20) $v = v_0$, we will obtain the parametric equations of the coordinate line:

$$x = \varphi(u, v_0), \qquad y = \psi(u, v_0). \tag{22.22}$$

Here the variable u appears as a parameter.

By assigning to v various possible constant values we will obtain a family of coordinate lines (v = const) in the xy-plane. In a similar manner, we obtain another family of coordinate lines (u = const).

If there is a one-to-one correspondence between domains D^* and D, the coordinate lines of a family never intersect, and through any one point of D passes only one line of each family.

The network of curvilinear coordinate lines in the xy-plane is a representation of a rectangular network in the uv-plane (see Fig. 22.13*a*, *b*).

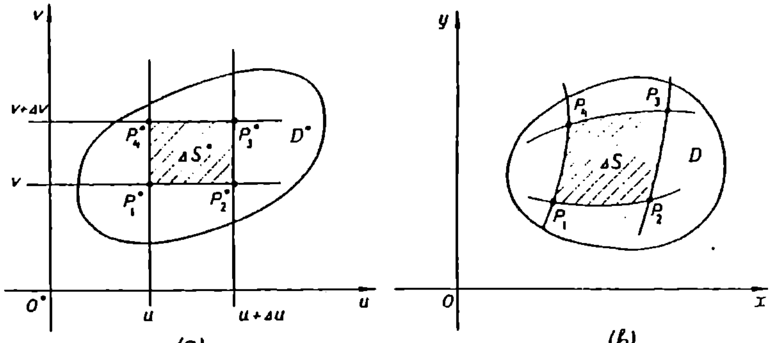


Fig. 22.14

Area element in curvilinear coordinates. Jacobian and its geometrical meaning. In a domain D in the uv-plane we draw a small rectangle $P_1^*P_2^*P_3^*P_4^*$ with sides parallel to the u- and v-coordinate axes and having lengths Δu and Δv, respectively (Fig. 22.14*a*). Its area is

$$\Delta S^* = \Delta u\, \Delta v. \tag{22.23}$$

For definiteness, we assume that $\Delta u > 0$ and $\Delta v > 0$.

The rectangle $P_1^*P_2^*P_3^*P_4^*$ is mapped into the curvilinear quadrangle $P_1P_2P_3P_4$ in D (Fig. 22.14*b*). If the vertices P_i^* ($i = 1, 2, 3, 4$) have the coordinates $P_1^*(u, v)$, $P_2^*(u + \Delta u, v)$, $P_3^*(u + \Delta u, v + \Delta v)$, $P_4^*(u, v + \Delta v)$, then by (22.20) the corresponding vertices P_i will have the coordinates

$$P_1(\varphi(u, v), \psi(u, v)), \qquad P_2(\varphi(u + \Delta u, v), \psi(u + \Delta u, v))$$
$$P_3(\varphi(u + \Delta u, v + \Delta v), \psi(u + \Delta u, v + \Delta v)),$$
$$P_4(\varphi(u, v + \Delta v), \psi(u, v + \Delta v)).$$

Using the Taylor formula for a function of two variables and keeping only first-order terms in Δu and Δv, we will determine approximately for $P_1P_2P_3P_4$ the coordinates of the vertices

$$\begin{aligned}
&P_1(\varphi,\ \psi) \qquad \text{(accurate value)},\\
&P_2\left(\varphi + \frac{\partial\varphi}{\partial u}\Delta u,\ \psi + \frac{\partial\psi}{\partial u}\Delta u\right),\\
&P_3\left(\varphi + \frac{\partial\varphi}{\partial u}\Delta u + \frac{\partial\varphi}{\partial v}\Delta v,\ \psi + \frac{\partial\psi}{\partial u}\Delta u + \frac{\partial\psi}{\partial v}\Delta v\right),\\
&P_4\left(\varphi + \frac{\partial\varphi}{\partial v}\Delta v,\ \psi + \frac{\partial\psi}{\partial v}\Delta v\right),
\end{aligned} \tag{22.24}$$

where φ, ψ and all their derivatives are computed at the point $(u,\ v)$.

Coordinates (22.24) indicate that, up to higher-order terms, the quadrangle $P_1P_2P_3P_4$ is a parallelogram. This follows from the fact that

$$\overrightarrow{P_1P_2} = \overrightarrow{P_4P_3} = \frac{\partial\varphi}{\partial u}\Delta u\,\mathbf{i} + \frac{\partial\psi}{\partial u}\Delta u\,\mathbf{j}.$$

The vector $\overrightarrow{P_1P_4}$ can be represented as

$$\overrightarrow{P_1P_4} = \frac{\partial\varphi}{\partial v}\Delta v\,\mathbf{i} + \frac{\partial\psi}{\partial v}\Delta v\,\mathbf{j}.$$

We will then be able to represent approximately the area ΔS of the quadrangle $P_1P_2P_3P_4$ through the length of the vector product $[\overrightarrow{P_1P_2},\ \overrightarrow{P_1P_4}]$, i.e.,

$$\Delta S \approx |[\overrightarrow{P_1P_2},\ \overrightarrow{P_1P_4}]| = \left\|\begin{matrix} \dfrac{\partial\varphi}{\partial u} & \dfrac{\partial\psi}{\partial u} \\ \dfrac{\partial\varphi}{\partial v} & \dfrac{\partial\psi}{\partial v} \end{matrix}\right\| \Delta u\,\Delta v.$$

The determinant

$$\begin{vmatrix} \dfrac{\partial\varphi(u,\ v)}{\partial u} & \dfrac{\partial\psi(u,\ v)}{\partial u} \\ \dfrac{\partial\varphi(u,\ v)}{\partial v} & \dfrac{\partial\psi(u,\ v)}{\partial v} \end{vmatrix} = J$$

is called the *functional determinant* of the functions $\varphi(u,\ v)$, $\psi(u,\ v)$, or the *Jacobian*.

Thus,

$$\Delta S \approx |J|\,\Delta u\,\Delta v. \tag{22.25}$$

The expression on the right of (22.25) is called an *area element in curvilinear coordinates*. Since $\Delta u\,\Delta v = \Delta S^*$, we obtain from (22.25)

$$\frac{\Delta S}{\Delta S^*} \approx |J|. \tag{22.26}$$

Equality (22.26) is approximate. In the limit, however, when the diameters of the elements ΔS^* and ΔS tend to zero, it becomes exact, i.e.,

$$|J(u, v)| = \lim_{\operatorname{diam} \Delta S^* \to 0} \frac{\Delta S}{\Delta S^*}. \tag{22.27}$$

It is seen from (22.26) and (22.27) that the absolute value of the Jacobian is a sort of local coefficient of 'extension" of the domain D^* (at a given point (u, v)) when mapped into D using the transformation formulas (22.20): $x = \varphi(u, v)$, $y = \psi(u, v)$.

Change of variables in the double integral. Suppose that two continuous functions $x = \varphi(u, v)$, $y = \psi(u, v)$ effect a one-to-one mapping of a domain D into D^* and have continuous partial derivatives of the first order. Suppose then that in D in the xy-plane a continuous function $z = f(x, y)$ is defined.

Corresponding to each value of the function $z = f(x, y)$ in D is an equal value of the function $z = F(u, v)$ in D^*, where $F(u, v) = f[\varphi(u, v), \psi(u, v)]$.

Consider integral sums for $z = f(x, y)$ in D and D^*. Making use of the fact that integral sums are set up in an arbitrary manner, we will form them so that they contain equal values of the function in D and D^*. We will get

$$\sum_D f(x, y)\Delta S \approx \sum_{D^*} F(u, v)|J|\,\Delta S^*, \tag{22.28}$$

where $\Delta S \approx |J|\Delta S^*$ and $J(u, v)$ is the Jacobian of $\varphi(u, v)$ and $\psi(u, v)$. Passing in (22.28) to the limit as the largest diameter d^* of the partial domains D_k^* tends to zero, we will have

$$\iint_D f(x, y)dS = \iint_{D^*} F(u, v)|J(u, v)|\,dS^*$$

or

$$\iint_D f(x, y)dx\,dy = \iint_{D^*} f[\varphi(u, v), \psi(u, v)]|J(u, v)|\;du\,dv, \tag{22.29}$$

where

$$J(u, v) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}.$$

Formula (22.29) is known as the *Ostrogradsky formula* to transform coordinates in a double integral.

Notice that as $d^* \to 0$ the largest of d of the partial domains in D will also tend to zero, because (22.20) is continuous.

The condition $J \neq 0$ is the condition that the mapping through the functions $x = \varphi(u, v)$, $y = \psi(u, v)$ be a locally one-to-one mapping.

Variables in a double integral are changed according to the following rules:

Theorem 22.4. *To transform a double integral specified in rectangular coordinates into a double integral in curvilinear coordinates we will have to replace in the integrand $f(x, y)$ the variables x and y by $\varphi(u, v)$ and $\psi(u, v)$ respectively, and the area element $dx\,dy$ by its expression in curvilinear coordinates: $dx\,dy = |J|\,du\,dv$.*

Example. Find the area of a figure bounded by the hyperbolas $xy = a^2$, $xy = b^2$, where $x > 0$, $y > 0$, $0 < a < b$, and by the straight lines $y = \alpha x$, $y = \beta x$, where $0 < \alpha < \beta$ (Fig. 22.15a).

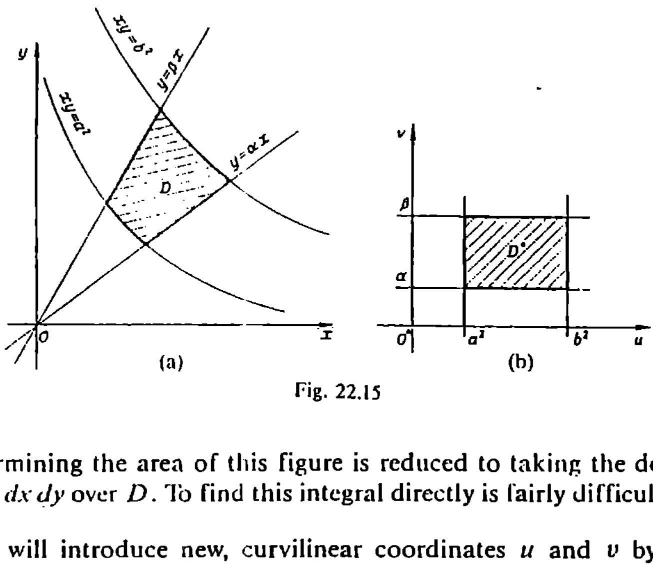


Fig. 22.15

◄ Determining the area of this figure is reduced to taking the double integral $\iint\limits_D dx\,dy$ over D. To find this integral directly is fairly difficult. Therefore, we will introduce new, curvilinear coordinates u and v by

$$xy = u \quad \text{and} \quad \frac{y}{x} = v. \tag{22.30}$$

It is clear then from the conditions of the problem that $a^2 \leqslant u \leqslant b^2$, $\alpha \leqslant v \leqslant \beta$. It follows that in the uv-plane we obtained the rectangle (Fig. 22.15b)

$$D^* = \{a^2 \leqslant u \leqslant b^2, \quad \alpha \leqslant v \leqslant \beta\}.$$

This is a simpler figure than the original one in D.

We now express x and y using (22.30) through u and v:

$$x = \sqrt{\frac{u}{v}}, \qquad y = \sqrt{uv}\,,$$

Then

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2\sqrt{uv}} & -\dfrac{\sqrt{u}}{2v\sqrt{v}} \\ \dfrac{1}{2}\sqrt{\dfrac{v}{u}} & \dfrac{1}{2}\sqrt{\dfrac{u}{v}} \end{vmatrix} = \frac{1}{2v}.$$

From (22.29) we obtain at $f(x, y) \equiv 1$

$$S = \iint\limits_{D} dx\,dy = \iint\limits_{D^*} |J|\,du\,dv = \int\limits_{a^2}^{b^2} du \int\limits_{\alpha}^{\beta} \frac{dv}{2v}$$

$$= u\,\Big|_{a^2}^{b^2} \cdot \frac{1}{2} \ln v\,\Big|_{\alpha}^{\beta} = \frac{b^2 - a^2}{2} \ln \frac{\beta}{\alpha}. \quad ►$$

Double integral in polar coordinates. Taking a double integral is frequently simplified by a change of rectangular coordinates x and y to polar coordinates ϱ and φ using the formulas

$$x = \varrho \cos \varphi, \qquad y = \varrho \sin \varphi, \qquad \text{where} \quad \varrho \geqslant 0, \quad 0 \leqslant \varphi \leqslant 2\pi. \tag{22.31}$$

Then

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial \varrho} & \dfrac{\partial x}{\partial \varphi} \\ \dfrac{\partial y}{\partial \varrho} & \dfrac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos \varphi & -\varrho \sin \varphi \\ \sin \varphi & \varrho \cos \varphi \end{vmatrix} = \varrho.$$

Considering that $|J| = \varrho$, we will derive the area element in polar coordinates

$$dS = \varrho\, d\varrho\, d\varphi \tag{22.32}$$

and the formula to pass over from an integral in rectangular coordinates to an integral in polar coordinates

$$\iint\limits_{D} f(x, y)\,dx\,dy = \iint\limits_{D^*} f[\varrho \cos \varphi, \varrho \sin \varphi]\varrho\, d\varrho\, d\varphi. \tag{22.33}$$

The area element in polar coordinates can be obtained from the following geometrical arguments (Fig. 22.16). We have

$$\Delta S = \text{area of } ODC - \text{area of } OAB$$

$$= \frac{1}{2}(\varrho + \Delta\varrho)^2 \Delta\varphi - \frac{1}{2}\varrho^2 \Delta\varphi \doteq \varrho\, \Delta\varrho\, \Delta\varphi + \frac{1}{2}(\Delta\varrho)^2 \Delta\varphi.$$

Discarding the infinitesimal quantity of a higher order, we will obtain

$$\Delta S = \varrho\, \Delta\varrho\, \Delta\varphi$$

and we take the area element in polar coordinates to be

$$dS = \varrho\, d\varrho\, d\varphi.$$

Consequently, to transform a double integral in rectangular coordinates into a double integral in polar coordinates we must in the integrand replace x and y by $r \cos\varphi$ and $\varrho \sin\varphi$, respectively, and replace the area element in rectangular coordinates $dx\,dy$ by an element in polar coordinates $\varrho\, d\varrho\, d\varphi$.

We now proceed to take the double integral in polar coordinates. As in the case of rectangular coordinates, we here also reduce the integral to the interated one.

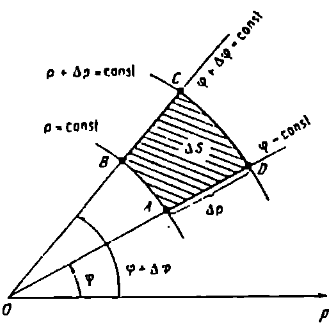


Fig. 22.16

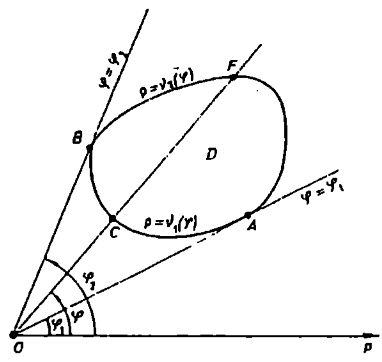


Fig. 22.17

Consider first the case where the pole O lies beyond the given domain D. Let D be such that any ray originating from the pole (the coordinate line $\varphi = \text{const}$) intersects its boundary at no more than two points or along a segment. Note the extreme values φ_1 and φ_2 of the polar angle φ, i.e., the polar angle φ varies from φ_1 to φ_2: $\varphi_1 \leqslant \varphi \leqslant \varphi_2$. Here φ_1 and φ_2 are the limits of external integration.

The ray $\varphi = \varphi_1$ passes through a point A of the contour of the domain D, and the ray $\varphi = \varphi_2$ through a point B (Fig. 22.17).

Points A and B divide the contour into two parts: ABC and AFB. Let their polar equations be $\varrho = \nu_1(\varphi)$ and $\varrho = \nu_2(\varphi)$, respectively, where $\nu_1(\varphi)$ and $\nu_2(\varphi)$ are single-valued continuous functions of φ satisfying the condition $\nu_1(\varphi) \leqslant \nu_2(\varphi)$ for all $\varphi \in (\varphi_1, \varphi_2)$.

The functions $\nu_1(\varphi)$ and $\nu_2(\varphi)$ are the limits of internal integration. Going over to iterated integrals gives

$$\iint_D F(\varrho, \varphi)\varrho \, d\varrho \, d\varphi = \int_{\varphi_1}^{\varphi_2} d\varphi \int_{\nu_1(\varphi)}^{\nu_2(\varphi)} F(\varrho, \varphi)\varrho \, d\varrho. \tag{22.34}$$

Specifically, the area S of D at $F(\varrho, \varphi) \equiv 1$ will be

$$S = \int_{\varphi_1}^{\varphi_2} d\varphi \int_{\nu_1(\varphi)}^{\nu_2(\varphi)} \varrho \, d\varrho = \frac{1}{2}\int_{\varphi_1}^{\varphi_2} [\nu_2^2(\varphi) - \nu_1^2(\varphi)] d\varphi.$$

Now suppose that the pole O lies within D. And suppose also that for the pole D is a "stellar" domain (Fig. 22.18), i.e., any ray φ = const

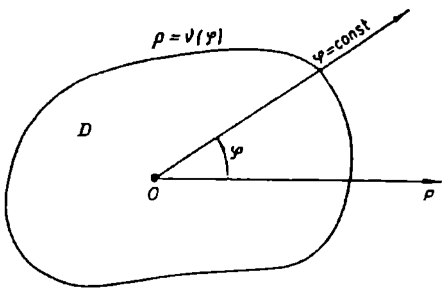


Fig. 22.18

cuts across the boundary of the domain only at one point or along a segment. Let $\varrho = \nu(\varphi)$ be the equation of the boundary in polar coordinates. Then,

$$\iint_D F(\varrho, \varphi)\varrho \, d\varrho \, d\varphi = \int_0^{2\pi} d\varphi \int_0^{\nu(\varphi)} F(\varrho, \varphi)\varrho \, d\varrho. \tag{22.35}$$

Example. Take the integral

$$L = \iint_D \frac{1}{\sqrt{1 + x^2 + y^2}} \, dx \, dy,$$

where D: $\{x^2 + y^2 \leqslant 1, x \geqslant 0, y \geqslant 0\}$ is the quarter of the unit circle in the first quadrant.

◄ Pass to polar coordinates

$$x = \varrho \cos\varphi, \qquad y = \varrho \sin\varphi.$$

Then the region of integration will be the rectangle

$$D^*: \left\{0 \leqslant \varrho \leqslant 1, \quad 0 \leqslant \varphi \leqslant \frac{\pi}{2}\right\}$$

and the integral can be transformed to yield

$$I = \iint_{D} \frac{\varrho\, d\varrho\, d\varphi}{\sqrt{1+\varrho^2}} = \int_0^{\pi/2} d\varphi \int_0^1 \frac{\varrho\, d\varrho}{\sqrt{1+\varrho^2}}$$

$$= \frac{\pi}{2}\sqrt{1+\varrho^2}\,\Big|_{\varrho=0}^{1} = \frac{\pi(\sqrt{2}-1)}{2}.$$

Remark. If the Jacobian is distinct from zero in D, then in a certain neighbourhood of each point in the domain we will have a one-to-one mapping.

It may , however, happen that the mapping of the entire domain is not one-to-one. Take, for example, the mapping given by the functions $x = e^u \cos v$, $y = e^u \sin v$, $-\infty < u, v < +\infty$. The Jacobian of these functions is

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} e^u \cos v & -e^u \sin v \\ e^u \sin v & e^u \cos v \end{vmatrix} = e^u$$

and hence is nonzero everywhere. Nevertheless, for $u = 0$, $v = 0$ and for $u = 0$, $v = 2\pi$ we obtain $x = 1$, $y = 0$, since the mapping is not one-to-one.

On the other hand, if at some point the Jacobian becomes zero, we may have a one-to-one mapping in the neighbourhood of this point. For example, for a mapping given by the functions $x = u^3$, $y = v^3$, $-\infty < u$, $v < +\infty$, the Jacobian $J = 9u^2v^2$ is zero at $u = 0$, or at $v = 0$, but the mapping is one-to-one.

The inverse mapping is given by the functions $u = \sqrt[3]{x}$, $v = \sqrt[3]{y}$, $-\infty < x, y < +\infty$, so that the point M $(x = 0, y = 0)$ is mapped into M^* $(u = 0, v = 0)$, and vice versa.

22.5 Surface Area. Surface Integral

Calculation of surface area. Consider a surface π that is uniquely projectable onto a domain D in the xy-plane. This means that this surface is described by the equation $z = f(x, y)$, where $P(x; y) \in D$.

We will suppose that the surface is *smooth*, i.e., that in D the function $f(x, y)$ is continuous and has continuous partial derivatives $f'_x(x; y)$ and $f'_y(x, y)$.

We will at first establish what we are going to call the area of a surface. We divide D into squarable subdomains $D_1, D_2, \ldots, D_n$ that do not contain common internal points, and will denote their areas by $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$, respectively. Let d be the largest of the diameters of the partial domains D_k ($k = 1, 2, \ldots, n$). In each D_k we select arbitrarily a point $P_k(\xi_k, \eta_k)$. The point P_k in π will have as its counterpart a point $M_k(\xi_k, \eta_k, \zeta_k)$, where $\zeta_k = f(\xi_k, \eta_k)$ (Fig. 22.19). We construct a plane tangent to π at M_k. Its equation is

$$z - \zeta_k = f'_x(\xi_k, \eta_k)(x - \xi_k) + f'_y(\xi_k, \eta_k)(y - \eta_k). \tag{22.36}$$

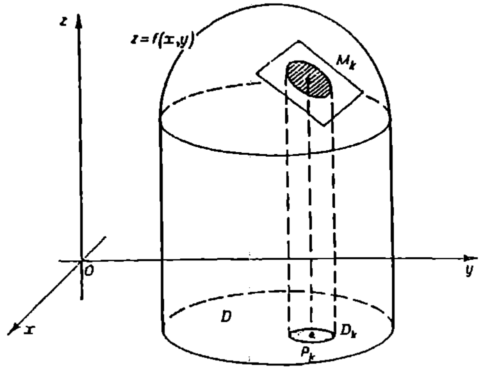


Fig. 22.19

We now construct on the boundary of D_k, as on a directrix, a cylindrical surface with a generatrix parallel to the z-axis. This surface will cut out a patch π_k with area $\Delta\sigma_k$ from the tangent surface through M_k. The patch π_k is projected on the xy-plane into D_k. Consider a sum

$$\sum_{k=1}^{n} \Delta\sigma_k. \tag{22.37}$$

Definition. If sum (22.37) has a definite limit S as $d \to 0$, i.e.,

$$\lim_{d\to 0} \sum_{k=1}^{n} \Delta\sigma_k = S, \tag{22.38}$$

then S is said to be the *surface area* of π. Here d is the largest of the diameters of D_k ($k = 1, 2, \ldots, n$).

Thus, we, as it were, replace the given surface by a scaly one, calculate the area of the scaly surface and pass to the limit as the diameter of scales tends to zero (as $d \to 0$ the diameters tend to zero as well).

We would now like to derive a formula to compute the surface area. It is well known that the area of the projection of a plane figure on some plane is equal to the product of the area of the figure being projected by the cosine of the acute angle between the projection plane and the plane in which the figure lies. We denote by γ_k the angle between the tangent plane to the surface π at M_k and the xy-plane (Fig. 22.20). Then

$$\Delta S_k = \Delta\sigma_k\,|\cos\gamma_k|,$$

hence

$$\Delta\sigma_k = \frac{\Delta S_k}{|\cos\gamma_k|}. \tag{22.39}$$

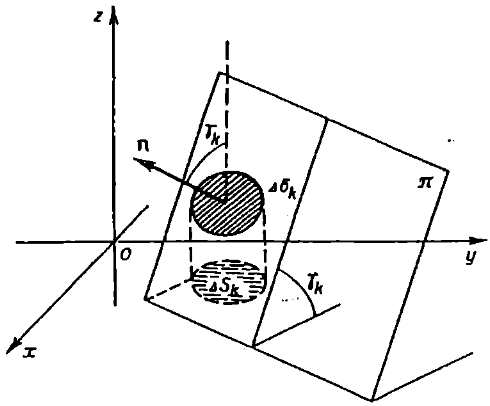


Fig. 22.20

But the angle γ_k is at the same time the angle between the z-axis and the normal to the tangent plane to (22.36). We denote by

$$\mathbf{n}_1 = \{f'_x(\xi_k, \eta_k),\ f'_y(\xi_k, \eta_k),\ -1\}$$

the normal to the plane tangent to the surface at M_k, and by $\mathbf{n}_2 = \{0, 0, 1\}$ the unit vector of the z-axis. We then obtain

$$\cos\gamma_k = \frac{|(\mathbf{n}_1, \mathbf{n}_2)|}{|\mathbf{n}_1||\mathbf{n}_2|} = \frac{1}{\sqrt{1 + [f'_x(\xi_k, \eta_k)]^2 + [f'_y(\xi_k, \eta_k)]^2}}.$$

Thus

$$\sum_{k=1}^{n} \Delta\sigma_k = \sum_{k=1}^{n} \sqrt{1 + [f'_x(\xi_k, \eta_k)]^2 + [f'_y(\xi_k, \eta_k)]^2}\,\Delta S_k. \tag{22.40}$$

As stated, $f_x'(x, y)$ and $f_y'(x, y)$ are continuous in D, and so is the function

$$\sqrt{1 + [f_x'(x, y)]^2 + [f_y'(x, y)]^2},$$

which is also integrable over D. Therefore, as $d \to 0$ the sum (22.40) has the finite limit

$$\lim_{d \to 0} \sum_{k=1}^{n} \Delta\sigma_k = \iint_D \sqrt{1 + [f_x'(x, y)]^2 + [f_y'(x, y)]^2}\, ds.$$

It follows from relation (22.38) defining the area S of π, that

$$S = \iint_{D_{xy}} \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy, \tag{22.41}$$

where D is the projection of π on the xy-plane.

The expression

$$d\sigma = \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy \tag{22.42}$$

is called a *surface element*.

If we project π on the xz-plane, we obtain

$$S = \iint_{D_{xz}} \sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2 + \left(\frac{\partial y}{\partial z}\right)^2}\, dx\, dz, \tag{22.43}$$

where D_{xz} is the projection of some regions of π on the xz-plane, and projecting π on the yz-plane we will have

$$S = \iint_{D_{yz}} \sqrt{1 + \left(\frac{\partial x}{\partial y}\right)^2 + \left(\frac{\partial x}{\partial z}\right)^2}\, dy\, dz, \tag{22.44}$$

where D_{yz} is the projection of the region on the yz-plane.

Example. Find the surface area of the sphere of radius R given by $x^2 + y^2 + z^2 = R^2$ with the centre at the origin of coordinates.

◄ The equation of the upper hemisphere is $z = \sqrt{R^2 - x^2 - y^2}$. Therefore,

$$\frac{\partial z}{\partial x} = -\frac{x}{\sqrt{R^2 - x^2 - y^2}}, \qquad \frac{\partial z}{\partial y} = -\frac{y}{\sqrt{R^2 - x^2 - y^2}}.$$

Hence

$$d\sigma = \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy$$

$$= \sqrt{1 + \frac{x^2 + y^2}{R^2 - x^2 - y^2}}\, dx\, dy = \frac{R\, dx\, dy}{\sqrt{R^2 - x^2 - y^2}}.$$

The domain of integration is D: $\{x^2 + y^2 \leqslant R^2\}$. And so the area will be

$$S = 2\iint_D \frac{R}{\sqrt{R^2 - x^2 - y^2}}\, dx\, dy = \left| \begin{array}{l} x = \varrho \cos\varphi \\ y = \varrho \sin\varphi \\ J = \varrho \end{array} \right|$$

$$= 2R\iint_{D^*} \frac{\varrho\, d\varrho\, d\varphi}{\sqrt{R^2 - \varrho^2}} = 2R\int_0^{2\pi} d\varphi \int_0^R \frac{\varrho\, d\varrho}{\sqrt{R^2 - \varrho^2}}$$

$$= 4\pi R\left(-\sqrt{R^2 - \varrho^2}\right)\Big|_{\varrho=0}^{R} = 4\pi R^2. \blacktriangleright$$

Note the following useful formulas:

(1) The surface element of the cylindrical surface of radius R is

$$d\sigma = R\, d\varphi\, dz. \tag{22.45}$$

(2) The surface element of the spherical surface of radius R is

$$d\sigma = R^2 \sin\theta\, d\theta\, d\varphi. \tag{22.46}$$

Using formula (22.46) for the surface element of a spherical surface, we derive the surface area of the sphere

$$S = 2\iint_\pi R^2 \sin\theta\, d\theta\, d\varphi = 2R^2 \int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin\theta\, d\theta$$

$$= 2R^2\, 2\pi(-\cos\theta)\Big|_{\theta=0}^{\pi/2} = 4\pi R^2.$$

Integral over a surface (surface integral of the first kind). Consider a continuous function $f(M)$ defined on a smooth surface π. Divide π into parts $\pi_1, \pi_2, \ldots, \pi_n$ with areas $\Delta\sigma_1, \Delta\sigma_2, \ldots, \Delta\sigma_n$, respectively; in each of the partial surfaces we take an arbitrary point $M_1, M_2, \ldots, M_n$ and form the sum

$$\sigma = \sum_{k=1}^{n} f(M_k)\Delta\sigma_k, \tag{22.47}$$

which we will call the *integral sum* for $f(M)$ over π.

Definition. If the largest of the diameters of π_k tends to zero and the integral sum (22.47) has a finite limit which is independent of the ways π is divided into parts or points M_k are chosen, then this limit is called the *integral of $f(M)$ over the surface area π (surface integral of the first kind)* and denoted as $\iint\limits_{\pi} f(M)d\sigma$ or $\iint\limits_{\pi} f(x, y, z)d\sigma$, where $d\sigma$ is the surface element.

The general properties of double integrals are valid for surface integrals. In particular, if π is broken into nonoverlapping parts $\pi_1, \pi_2, \ldots, \pi_n$, then

$$\iint\limits_{\pi} f(M)d\sigma = \sum_{k=1}^{n} \iint\limits_{\pi_k} f(M)d\sigma. \tag{22.48}$$

Theorem 22.5. *If π is a smooth surface given by the equation $z = \varphi(x, y)$ and $\varphi(x, y)$ has continuous partial derivatives in a certain domain D_1, and if $f(x, y, z)$ is a continuous function defined in π, then for any closed bounded domain $D \subset D_1$ there holds the identity*

$$\iint\limits_{\pi} f(x, y, z)d\sigma = \iint\limits_{D} f(x, y, \varphi(x, y))\sqrt{1 + (\varphi_x')^2 + (\varphi_y')^2}\, dx\, dy. \tag{22.49}$$

The surface integral on the left exists if there exists the double integral on the right of (22.49).

The integral $\iint\limits_{\pi} \mu(P)d\sigma$, where $\mu(P) \geqslant 0$ in π, can be treated as the mass m of the shell represented by the surface π, over which the mass is distributed with the surface density $\mu = \mu(P)$.

Example. Find the mass of the parabolic shell

$$z = \frac{1}{2}(x^2 + y^2), \qquad 0 \leqslant z \leqslant 1,$$

whose density varies according to the law $\mu = z$ (Fig. 22.21).

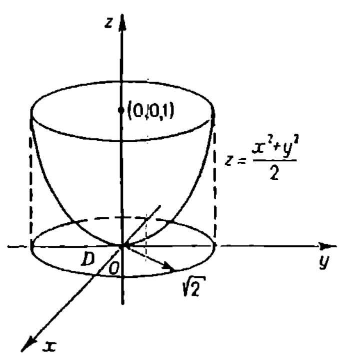


Fig. 22.21

◄ We have

$$m = \iint_{\pi} \mu(P)d\sigma = \iint_{D} z(x, y)\sqrt{1 + (z_x')^2 + (z_y')^2}\, dx\, dy$$

$$= \frac{1}{2}\iint_{D} (x^2 + y^2)\sqrt{1 + x^2 + y^2}\, dx\, dy$$

$$= \frac{1}{2}\int_0^{2\pi} d\varphi \int_0^{\sqrt{2}} \varrho^2\sqrt{1 + \varrho^2}\, \varrho\, d\varrho = \pi\int_0^{\sqrt{2}} \sqrt{1 + \varrho^2}\, \varrho^3\, d\varrho$$

$$= \left| \begin{array}{lll} \sqrt{1 + \varrho^2} = t, & \varrho = \sqrt{t^2 - 1}, & d\varrho = \dfrac{t\, dt}{\sqrt{t^2 - 1}} \\ \varrho_1 = 0, & t_1 = 1, & \\ \varrho_2 = \sqrt{2}, & t_2 = \sqrt{3} & \end{array} \right|$$

$$= \pi\int_1^{\sqrt{3}} (\sqrt{t^2 - 1})^3 \frac{t^2\, dt}{\sqrt{t^2 - 1}} = \pi\int_1^{\sqrt{3}} (t^2 - 1)t^2\, dt$$

$$= \pi\left(\frac{t^5}{5} - \frac{t^3}{3}\right)\Bigg|_1^{\sqrt{3}} = \frac{2\pi}{15}(6\sqrt{3} + 1). \quad ►$$

22.6 Triple Integrals

Suppose that we are given a material body occupying a three-dimensional region Ω filled with matter. Suppose also that at each point we know the density

$$\mu = \mu(P) = \mu(x, y, z).$$

It is required to find the mass m of the body.

We will break Ω into nonoverlaping cubable (i.e., having a volume) parts $\Omega_1, \Omega_2, \ldots, \Omega_n$ having volumes $\Delta V_1, \Delta V_2, \ldots, \Delta V_n$, respectively. In each of Ω_k we select an arbitrary point P_k. We assume approximately that throughout Ω_k the density is constant and equal to $\mu(P_k)$. The mass Δm_k of this part of the body will then be given by the approximate relation

$$\Delta m_k \simeq \mu(P_k)\Delta V_k.$$

The mass of the entire body will be

$$m \simeq \sum_{k=1}^{n} \mu(P_k)\Delta V_k. \tag{22.50}$$

We denote by d the largest of the diameters of Ω_k ($k = 1, 2, \ldots, n$). If (22.50) has a finite limit as $d \to 0$, which is independent of the way Ω has been divided into partial regions or $P_k \in \Omega_k$ have been selected, then

this limit is taken to be the mass of the body

$$m = \lim_{d \to 0} \sum_{k=1}^{n} \mu(P_k)\Delta V_k. \tag{22.51}$$

On the other hand, this limit is known as the ***triple integral*** of the function $\mu(P)$ over the domain Ω and is denoted by

$$\iiint_{\Omega} \mu(P)dv = \lim_{d \to 0} \sum_{k=1}^{n} \mu(P_k)\,\Delta V_k.$$

Hence

$$m = \iiint_{\Omega} \mu(P)dv = \iiint_{\Omega} \mu(x, y, z)dx\,dy\,dz. \tag{22.52}$$

Here $dx\,dy\,dz$ is an element of volume dV in rectilinear coordinates.

We will now give the formal definition of the triple integral. Consider a closed cubable domain Ω in which a bounded function $f(P)$, $P \in \Omega$, is defined. We divide Ω into n nonoverlapping cubable parts $\Omega_1, \Omega_2, \ldots, \Omega_n$, and denote their volumes by $\Delta V_1, \Delta V_2, \ldots, \Delta V_n$, respectively. In each Ω_k we arbitrarily choose a point $P_k(x_k, y_k, z_k)$ and form the integral sum

$$\sigma = \sum_{k=1}^{n} f(P_k)\Delta V_k.$$

Let d be the largest of the diameters of Ω_k $(k = 1, 2, \ldots, n)$.

Definition. If when $d \to 0$ the integral sums σ have a limit that is dependent neither on the way in which Ω is broken down into Ω_k nor on the choice of points $P_k \in \Omega_k$, then the limit is called the *triple integral* of $f(x, y, z)$ over Ω and is denoted by

$$\iiint_{\Omega} f(x, y, z)dv \qquad \text{or} \qquad \iiint_{\Omega} f(P)dv.$$

The function $f(x, y, z)$ is then called *integrable in* Ω. By definition, we thus have

$$\iiint_{\Omega} f(x, y, z)dv = \lim_{d \to 0} \sum_{k=1}^{n} f(x_k, y_k, z_k)\Delta V_k. \tag{22.53}$$

Theorem 22.6. *If a function $f(x, y, z)$ is continuous in a closed cubable domain Ω, then it is integrable in that domain.*

Properties of triple integrals. Triple integrals are similar in their behaviour to double integrals. Therefore, we can easily list them.

Let the functions $f(P)$ and $\varphi(P)$ be integrable in a cubable domain Ω.

(1) *Linearity:*

$$\iiint_{\Omega} (\alpha f(P) + \beta\varphi(P))dv = \alpha \iiint_{\Omega} f(P)dv + \beta \iiint_{\Omega} \varphi(P)dv,$$

where α and β are arbitrary real constants.

(2) If $f(P) \leqslant \varphi(P)$ everywhere in Ω, then

$$\iiint_{\Omega} f(P)dv \leqslant \iiint_{\Omega} \varphi(P)dv.$$

(3) If $f(P) \equiv 1$ in Ω, then

$$\iiint_{\Omega} dv = V,$$

where V is the volume of Ω.

(4) If a function $f(P)$ is continuous in a closed cubable domain Ω and M and m are its largest and smallest values, respectively, in Ω, then

$$mV \leqslant \iiint_{\Omega} f(P)dv \leqslant MV,$$

where V is the volume of Ω.

(5) *Additivity.* If a region Ω is broken into cubable domains Ω_1 and Ω_2 without common internal points and $f(P)$ is integrable in Ω, then $f(P)$ is integrable in each of Ω_1 and Ω_2, and

$$\iiint_{\Omega} f(P)dv = \iiint_{\Omega_1} f(P)dv + \iiint_{\Omega_2} f(P)dv.$$

(6) **Theorem 22.7** (mean value theorem). *If a function $f(P)$ is continuous in a closed cubable domain Ω, then there exists a point $P_m \in \Omega$ such that*

$$\iiint_{\Omega} f(P)dv = f(P_m)V,$$

where V is the volume of Ω.

22.7 Taking Triple Integral in Rectangular Coordinates

As with double integrals, the problem comes down to taking iterated integrals.

Suppose that $f(x, y, z)$ is a continuous function in a domain Ω. Then there exist all the integrals described below.

We will first take the case where Ω is a rectangular parallelepiped

$$\Omega: \{a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d, \quad l \leqslant z \leqslant m\},$$

whose projection on the yz-plane is a rectangle R

$$R: \{c \leqslant y \leqslant d, \quad l \leqslant z \leqslant m\}.$$

We will then have

$$\iiint_{\Omega} f(x, y, z)dv = \int_a^b dx \iint_R f(x, y, z)ds. \tag{22.54}$$

Substituting the iterated integral for the double integral gives

$$\iiint_{\Omega} f(x, y, z)dv = \int_a^b dx \int_c^d dy \int_l^m f(x, y, z)dz. \tag{22.55}$$

When Ω is a rectangular parallelepiped, we could thus reduce the taking of the triple integral to taking three conventional integrals in a row.

Formula (22.55) can be rewritten to give

$$\iiint_{\Omega} f(x, y, z)dv = \iint_D \left(\int_l^m f(x, y, z)dz\right)dx\,dy, \tag{22.56}$$

where D: $\{a \leqslant x \leqslant b, c \leqslant y \leqslant d\}$ is a rectangle in the xy-plane. This rectangle is an orthogonal projection of the parallelepiped on the xy-plane.

Let us now consider a domain Ω such that a surface S that bounds Ω is intersected at no more than two points by any straight line parallel to the z-axis (Fig. 22.22). Suppose that the surface S_1 that bounds Ω from below is described by the equation $z = \varphi_1(x, y)$, and the surface S_2 that bounds Ω from above is described by the equation $z = \varphi_2(x, y)$.

We project S_1 and S_2 on the xy-plane into a domain D bounded by a curve L. The body Ω is bounded by a cylindrical surface with a generatrix parallel to the z-axis and with the curve L, which here acts like a directrix.

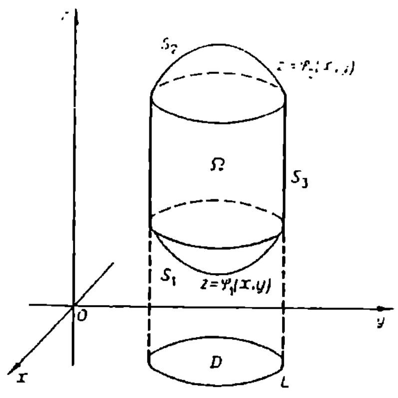


Fig. 22.22

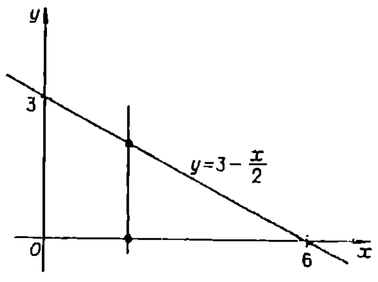


Fig. 22.23

In analogy with (22.56) we will then obtain

$$\iiint\limits_{\Omega} f(x,\ y,\ z)dv = \iint\limits_{D} \left[\int\limits_{\varphi_1(x,y)}^{\varphi_2(x,y)} f(x,\ y,\ z)dz\right] dx\,dy. \tag{22.57}$$

If the domain D in the xy-plane is a curvilinear trapezoid bounded by the curves $y = \psi_1(x)$ and $y = \psi_2(x)$ $(a \leqslant x \leqslant b)$, then the double integral in (22.57) can be reduced to an iterated integral, and so we will arrive at

$$\iiint\limits_{\Omega} f(x,\ y,\ z)dv = \int\limits_a^b dx \int\limits_{\psi_1 x}^{\psi_2 x} dy \int\limits_{\varphi_1(x,y)}^{\varphi_2(x,y)} f(x,\ y,\ z)dz. \tag{22.58}$$

This relation is a generalization of formula (22.55).

Example. Calculate the volume of the tetrahedron bounded by the planes $x = 0$, $y = 0$, $z = 0$ and $x + 2y + z - 6 = 0$.

◄ The projection of the tetrahedron on the xy-plane is the triangle formed by the straight lines $x = 0$, $y = 0$ and $x + 2y = 6$, so that x varies from 0 to 6, and at a fixed x $(0 \leqslant x \leqslant 6)$ y varies from 0 to $3 - x/2$ (Fig. 22.23). If x and y are fixed, the point can move along the vertical from the plane $z = 0$ to the plane $x + 2y + z - 6 = 0$, i.e., z varies from 0 to $6 - x - 2y$.

From (22.58) we obtain at $f(x,\ y,\ z) \equiv 1$

$$V = \int\limits_0^6 dx \int\limits_0^{3-\frac{x}{2}} dy \int\limits_0^{6-x-2y} dz = \int\limits_0^6 dx \int\limits_0^{3-\frac{x}{2}} (6 - x - 2y)dy$$

$$= \int\limits_0^6 [(6 - x)y - y^2]\, dx \,\Big|_{y=0}^{y=3-\frac{x}{2}} = \frac{1}{2}\int\limits_0^6 (6 - x)^2 dx$$

$$= -\frac{1}{2}\int\limits_0^6 (6 - x)^2 d(6 - x) = -\frac{1}{2}\,\frac{(6 - x)^3}{3}\,\Big|_0^6 = 36. \ ►$$

22.8 Taking Triple Integral in Cylindrical and Spherical Coordinates

In a triple integral variables are changed in much the same manner as in a double integral. Suppose we have a function $f(x,\ y,\ z)$ that is continuous in a closed cubable domain Ω, and the functions

$$x = x(\xi,\ \eta,\ \tau), \qquad y = y(\xi,\ \eta,\ \tau), \qquad z = z(\xi,\ \eta,\ \tau) \tag{22.59}$$

are continuous together with their partial derivatives of the first order in a closed cubable domain Ω^*.

Suppose further that the function (22.59) establishes a one-to-one correspondence between all the points (ξ, η, τ) of Ω^* and all the points (x, y, z) of Ω.

Then a *change of variables in a triple integral* is governed by the following formula:

$$\iiint\limits_{\Omega} f(x, y, z)\,dx\,dy\,dz$$
$$= \iiint\limits_{\Omega^*} f[x(\xi, \eta, \tau), y(\xi, \eta, \tau), z(\xi, \eta, \tau)]\,|J|\,d\xi\,d\eta\,d\tau, \quad (22.60)$$

where J is the Jacobian of the set of functions (22.59):

$$J = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \tau} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \tau} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \tau} \end{vmatrix}.$$

In actual practice, in taking triple integrals rectangular coordinates are often changed to cylindrical and spherical ones.

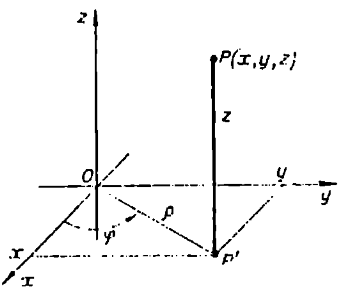


Fig. 22.24

Triple integrals in cylindrical coordinates. In cylindrical coordinates the position of a point P in space is determined by three numbers ϱ, φ, z, where ϱ and φ are the polar coordinates of the projection P' of P onto the xy-plane and z is the applicate of the point P. The numbers ϱ, φ, z are termed the *cylindrical coordinates* of P (Fig. 22.24).

Clearly,

$$0 \leqslant \varrho < +\infty,$$
$$0 \leqslant \varphi < 2\pi,$$
$$-\infty < z < +\infty.$$

In cylindrical coordinates the coordinate surfaces ϱ = const, φ = const, and z = const are respectively circular cylinders with axes coincident with the z-axis; half-planes adjacent to the z-axis, and planes parallel to the xy-plane.

The relation of rectangular coordinates and cylindrical coordinates is apparent from the figure

$$\begin{cases} x = \varrho \cos \varphi, \\ y = \varrho \sin \varphi, \\ z = z. \end{cases} \tag{22.61}$$

For the set (22.61) that maps the domain Ω into Ω^* we have

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial \varrho} & \dfrac{\partial x}{\partial \varphi} & \dfrac{\partial x}{\partial z} \\ \dfrac{\partial y}{\partial \varrho} & \dfrac{\partial y}{\partial \varphi} & \dfrac{\partial y}{\partial z} \\ \dfrac{\partial z}{\partial \varrho} & \dfrac{\partial z}{\partial \varphi} & \dfrac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos \varphi & -\varrho \sin \varphi & 0 \\ \sin \varphi & \varrho \cos \varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = \varrho.$$

Since $\varrho \geqslant 0$, then $|J| = \varrho$ and for triple integrals the formula (22.60) for the transition from rectangular coordinates to cylindrical ones becomes

$$\iiint\limits_{\Omega} f(x, y, z)dx\,dy\,dz = \iiint\limits_{\Omega^*} f(\varrho \cos \varphi, \varrho \sin \varphi, z)\varrho\, d\varrho\, d\varphi\, dz. \tag{22.62}$$

The expression $dv = \varrho\, d\varrho\, d\varphi\, dz$ is called a *volume element* in cylindrical coordinates. It can be deduced from geometrical cosiderations. We break Ω into elementary volumes by the coordinate surfaces ϱ = const, φ = const, z = const. The elementary volumes will be curvilinear prisms (Fig. 22.25). We see that

$$\begin{aligned} \Delta V &= \frac{1}{2}(\varrho + \Delta\varrho)^2 \Delta\varphi\, \Delta z - \frac{1}{2}\varrho^2 \Delta\varphi\, \Delta z \\ &= \varrho\, \Delta\varrho\, \Delta\varphi\, \Delta z + \frac{1}{2}(\Delta\varrho)^2 \Delta\varphi\, \Delta z. \end{aligned}$$

We discard the infinitesimal quantity of a higher order to get

$$\Delta V = \varrho\, \Delta\varrho\, \Delta\varphi\, \Delta z.$$

We take a volume element in cylindrical coordinates to be

$$dv = \varrho\, d\varrho\, d\varphi\, dz.$$

Example. Find the volume of the body bounded by the surfaces $z = x^2 + y^2$ and $z = 2 - x^2 - y^2$ (Fig. 22.26).

◄ We pass over to cylindrical coordinates $x = \varrho\cos\varphi$, $y = \varrho\sin\varphi$, $z = z$. The surfaces will then be described by the equations $z = \varrho^2$ and $z = 2 - \varrho^2$. Obviously, they intersect along the line

$$L: \begin{cases} \varrho = 1 & \text{(cylinder)}, \\ z = 1 & \text{(plane)}, \end{cases}$$

whose projection on the xy-plane will be

$$\varrho = 1, \qquad z = 0.$$

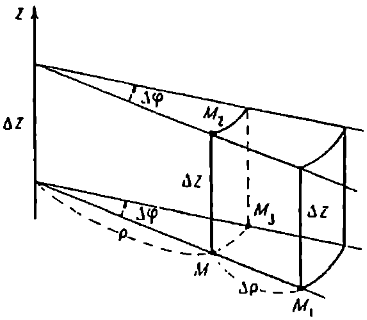


Fig. 22.25

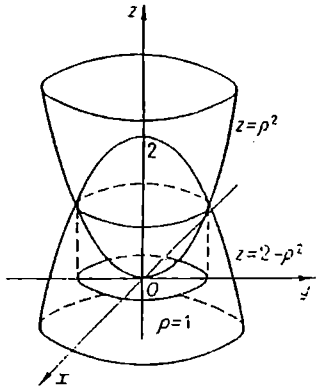


Fig. 22.26

Thus, $0 \leqslant \varrho \leqslant 1$, $0 \leqslant \varphi \leqslant 2\pi$, $\varrho^2 \leqslant z \leqslant 2 - \varrho^2$ and we will find the volume from (22.62), where we will set $f \equiv 1$.

$$V = \iiint\limits_{\Omega} dx\,dy\,dz = \iiint\limits_{\Omega^*} \varrho\,d\varrho\,d\varphi\,dz$$

$$= \int_0^{2\pi} d\varphi \int_0^1 \varrho\,d\varrho \int_{\varrho^2}^{2-\varrho^2} dz = 2\pi \int_0^1 (2 - 2\varrho^2)\varrho\,d\varrho$$

$$= 4\pi\left[\frac{\varrho^2}{2} - \frac{\varrho^4}{4}\right]\Bigg|_{\varrho=0}^{1} = \pi. \blacktriangleright$$

Triple integral in spherical coordinates. In spherical coordinates the position of $P(x, y, z)$ in space is determined by three numbers r, φ, θ, where r is the distance from the origin of coordinates to the point P; φ is the angle between the x-axis and the projection of the radius vector OP of

If $\gamma = \text{const}$, then $m = \gamma S$, where S is the area of D, and formulas (22.67) become

$$x_c = \frac{\iint\limits_D x\,dx\,dy}{S}, \quad y_c = \frac{\iint\limits_D y\,dx\,dy}{S}. \tag{22.68}$$

Example. Find the centre of mass of a uniform plane figure bounded by $y = \cos x$, $0 \leqslant x \leqslant \pi/2$, and the x- and y-axes.

◄ The figure being uniform, we will seek the centre of mass using the formulas (22.68).

To begin with, we will find the area S of the figure

$$S = \int\limits_0^{\pi/2} \cos x\,dx = \sin x \Big|_0^{\frac{\pi}{2}} = 1,$$

and then the static moments M_x and M_y

$$M_x = \iint\limits_D y\,dx\,dy = \int\limits_0^{\pi/2} dx \int\limits_0^{\cos x} y\,dy = \int\limits_0^{\pi/2} \frac{y^2}{2}\Big|_{y=0}^{\cos x} dx = \frac{1}{2}\int\limits_0^{\pi/2} \cos^2 x\,dx$$

$$= \frac{1}{4}\int\limits_0^{\pi/2} (1 + \cos 2x)\,dx = \frac{1}{4}\left(x + \frac{\sin 2x}{2}\right)\Big|_0^{\frac{\pi}{2}} = \frac{\pi}{8}.$$

$$M_y = \iint\limits_D x\,dx\,dy = \int\limits_0^{\pi/2} x\,dx \int\limits_0^{\cos x} dy = \int\limits_0^{\pi/2} x\cos x\,dx$$

$$= (x\sin x + \cos x)\Big|_0^{\frac{\pi}{2}} = \frac{\pi}{2} - 1.$$

Now from (22.68) we get

$$x_c = \frac{M_y}{S} = \frac{\pi}{2} - 1, \quad y_c = \frac{M_x}{S} = \frac{\pi}{8}. \blacktriangleright$$

Moments of inertia of a plane figure relative to coordinate axes. Following the same arguments, we readily establish that elementary moments of inertia relative to the x- and y-axes will respectively be

$$dI_x = y^2\,dm = y^2\gamma(x, y)\,ds = y^2\gamma(x, y)\,dx\,dy,$$
$$dI_y = x^2\,dm = x^2\gamma(x, y)\,ds = x^2\gamma(x, y)\,dx\,dy.$$

Integrating over the plane figure D, we will arrive at the formulas for the

moments themselves

$$I_x = \iint_D y^2 \gamma(x, y)\, dx\, dy, \tag{22.69}$$

$$I_y = \iint_D x^2 \gamma(x, y)\, dx\, dy, \tag{22.70}$$

where, as before, $\gamma(x, y)$ is the surface density distribution of mass.

Calculation of the mass of a body. When discussing the problem yielding a triple integral we have shown that if we know the density distribution $\mu(x, y, z)$ at each point of a body Ω, then the mass of the body will be given by

$$m = \iiint_\Omega \mu(x, y, z)\, dx\, dy\, dz. \tag{22.71}$$

We suppose that the function $\mu(x, y, z)$ is continuous in Ω.

Example. Calculate the mass m of a body bounded by the hemispheres $z = \sqrt{a^2 - x^2 - y^2}$ and $z = \sqrt{b^2 - x^2 - y^2}$ $(a < b)$ and the xy-plane if the density at each point is proportional to the distance of the point from the origin of coordinates.

◄ As stated, the density μ at the point (x, y, z) is given by $\mu(x, y, z) = k\sqrt{x^2 + y^2 + z^2}$, where k is the coefficient of proportionality. Then

$$m = \iiint_\Omega k\sqrt{x^2 + y^2 + z^2}\, dx\, dy\, dz.$$

Changing to spherical coordinates gives

$$a \leqslant r \leqslant b, \qquad 0 \leqslant \varphi \leqslant 2\pi, \qquad 0 \leqslant \theta \leqslant \frac{\pi}{2}$$

and

$$m = k\int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin\theta\, d\theta \int_a^b r^3\, dr$$

$$= k\, 2\pi(-\cos\theta)\Big|_0^{\frac{\pi}{2}} \cdot \frac{r^4}{4}\Big|_a^b = \frac{k\pi}{2}(b^4 - a^4) \quad ►$$

Static moments of a body relative to coordinate planes. Centre of mass. Recall that we determined static moments and the centre of mass of a plane figure using double integrals; see (22.65-67).

Static moments of a body Ω relative to coordinate planes and the centre of mass of Ω are sought for using triple integrals. For example, the elementary static moment relative to the xy-plane will be

$$dK_{xy} = z\, dm = z\,\mu(x, y, z)\, dv = z\,\mu(x, y, z)\, dx\, dy\, dz,$$

where $\mu(x, y, z)$ is the density. The static moment is thus given by

$$K_{xy} = \iiint\limits_{\Omega} z\,\mu(x, y, z)\,dx\,dy\,dz. \tag{22.72}$$

Likewise, we write the static moments relative to the xz- and yz-planes

$$K_{xz} = \iiint\limits_{\Omega} y\,\mu(x, y, z)\,dx\,dy\,dz,$$

$$K_{yz} = \iiint\limits_{\Omega} x\,\mu(x, y, z)\,dx\,dy\,dz.$$

After having found the mass m and the static moments of the body Ω, we can readily determine the coordinates of the centre of mass of the body

$$x_c = \frac{\iiint\limits_{\Omega} x\,\mu(x, y, z)\,dx\,dy\,dz}{m},$$

$$y_c = \frac{\iiint\limits_{\Omega} y\,\mu(x, y, z)\,dx\,dy\,dz}{m}, \tag{22.73}$$

$$z_c = \frac{\iiint\limits_{\Omega} z\,\mu(x, y, z)\,dx\,dy\,dz}{m}.$$

If the body is uniform, then the density $\mu = \text{const}$ in the domain Ω, and so (22.73) become simpler, since in the numerator we can take $\mu = \text{const}$ outside the sign of integral and cancel out μ in the numerator and denominator (because $m = \mu V$). We finally have

$$x_c = \frac{\iiint\limits_{\Omega} x\,dx\,dy\,dz}{V}, \quad y_c = \frac{\iiint\limits_{\Omega} y\,dx\,dy\,dz}{V},$$

$$z_c = \frac{\iiint\limits_{\Omega} z\,dx\,dy\,dz}{V}, \tag{22.74}$$

where V is the volume of Ω.

Example. Find the coordinates of the centre of mass of a uniform hemisphere of radius R.

◄ We suppose that the centre of the sphere lies at the origin of coordinates, and that the hemisphere lies below the xy-plane. By symmetry we then have $x_c = 0$ and $y_c = 0$.

The volume of the hemisphere is

$$V = \frac{2}{3}\pi R^3.$$

We now find the static moment relative to the xy-plane

$$K_{xy} = \iiint\limits_{\Omega} z\, dx\, dy\, dz = \int\limits_0^{2\pi} d\varphi \int\limits_0^{\pi/2} \sin\theta \cos\theta\, d\theta \int\limits_0^{R} r^3 dr$$

$$= 2\pi \frac{\sin^2\theta}{2}\Bigg|_0^{\pi/2} \cdot \frac{r^4}{4}\Bigg|_0^R = \frac{\pi}{4} R^4.$$

Hence

$$z_c = \frac{K_{xy}}{V} = \frac{\pi R^4}{4} : \frac{2\pi R^3}{3} = \frac{3}{8} R.$$

The centre of mass is thus $C(0, 0, 3R/8)$. ►

22.10 Improper Multiple Integrals over Unbounded Domains

We now turn to the integration of a function of several variables over an unbounded domain D. We then choose a sequence of bounded domains of integration $D_1, D_2, D_3, \ldots$, which exhaust D in a monotone manner, i.e., $D_n \subset D_{n+1}\ \forall n$ and $D_n \to D$ as $n \to \infty$. For example, if the domain of integration D coincides with the entire xy-plane, then we may take the sequence $\{D_n\}$ to be the collection of circles $x^2 + y^2 \leqslant a_n^2$, $a_n < a_{n+1}$, $n = 1, 2, \ldots$, with centre at the origin of coordinates.

Definition. The improper integral of a function $f(x, y)$ over an unbounded domain D is the limit of the sequence of integrals

$$\lim_{n\to\infty} \iint\limits_{D_n} f(x, y)\, dx\, dy \tag{22.75}$$

independent of the choice of the sequence D_n.

Thus, by definition

$$\iint\limits_{D} f(x, y)\, dx\, dy = \lim_{n\to\infty} \iint\limits_{D_n} f(x, y)\, dx\, dy. \tag{22.76}$$

If the limit (22.75) exists and is finite, then the improper integral over an unbounded domain is said to be *convergent*, otherwise it is *divergent*.

Example. Examine for convergence the integral

$$\iint\limits_{D} \frac{dx\, dy}{(x^2 + y^2 + 1)^2}, \tag{22.77}$$

where the integration domain is $D: \{-\infty < x < +\infty,\ -\infty < y < +\infty\}$.

◄ Since the domain D coincides with the entire xy-plane, then we choose as a sequence of domains D_n the circles $x^2 + y^2 \leqslant n^2$ of radius n ($n = 1, 2, \ldots$).

Going over to polar coordinates gives

$$\iint_D \frac{dx\,dy}{(x^2+y^2+1)^2} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{dx\,dy}{(x^2+y^2+1)^2}$$

$$= \lim_{n\to\infty}\int_0^{2\pi} d\varphi \int_0^{n} \frac{\varrho\,d\varrho}{(\varrho^2+1)^2} = 2\pi \lim_{n\to\infty}\left(-\frac{1}{2(\varrho^2+1)}\right)\Bigg|_{\varrho=0}^{\varrho=n}$$

$$= 2\pi \lim_{n\to+\infty}\left(\frac{1}{2} - \frac{1}{2(1+n^2)}\right) = \pi.$$

Therefore, integral (22.77) converges. ►

An integral over an unbounded domain D must obey the following comparison test: if $0 \leqslant f(x, y) \leqslant g(x, y)\,\forall (x, y) \in D$ and the integral $\iint_D g(x, y)\,dx\,dy$ converges, then so does the integral $\iint_D f(x, y)\,dx\,dy$. Conversely, if the former integral diverges, then so does the latter one.

Integrals that are convergent over the entire plane can be taken using repeated integration

$$\iint_D f(x, y)\,dx\,dy = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y)\,dy = \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(x, y)\,dx. \tag{22.78}$$

Example. Take the integral $\int_{-\infty}^{+\infty} e^{-x^2}\,dx$. Пуассона

◄ Since $I = \int_{-\infty}^{+\infty} e^{-x^2}\,dx = \int_{-\infty}^{+\infty} e^{-y^2}\,dy$, then according to (22.78)

$$I^2 = \int_{-\infty}^{+\infty} e^{-x^2}\,dx \int_{-\infty}^{+\infty} e^{-y^2}\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(x^2+y^2)}\,dx\,dy.$$

If then in the double integral we pass over to polar coordinates, we obtain the domain G: $\{0 \leqslant \varphi \leqslant 2\pi,\ 0 \leqslant \varrho < +\infty\}$. Accordingly,

$$I^2 = \int_0^{2\pi} d\varphi \int_0^{+\infty} e^{-\varrho^2}\varrho\,d\varrho = 2\pi\left(-\frac{e^{-\varrho^2}}{2}\right)\Bigg|_0^{\infty} = \pi,$$

hence

$$I = \int_{-\infty}^{+\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$ ►

In much the same manner we find improper integrals of a function of three, four and more variables over unbounded domains.

Exercises

Take the following double integrals:

1. $\int_1^3 dy \int_2^5 x^2 y\,dx$. 2. $\int_3^4 dx \int_1^2 \frac{1}{(x+y)^2}\,dy$.

3. $\int_{b/2}^{b} \varrho\,d\varrho \int_0^{\pi/2} d\theta$. 4. $\int_0^1 dx \int_0^1 (x^2+y^2)\,dy$.

Change the order of integration. Preliminarily draw the domain of integration:

5. $\int_3^4 dy \int_1^2 f(x, y)\,dx$. 6. $\int_0^a dx \int_0^{\sqrt{a^2-x^2}} f(x, y)\,dy$. 7. $\int_0^3 dx \int_{x/3}^{2x} f(x, y)\,dy$.

8. $\int_{-6}^{2} dy \int_{y^2/4-1}^{2-y} f(x, y)\,dx$. 9. $\int_{-2}^{0} dy \int_{y^2-4}^{3} dx$. 10. $\int_0^1 dx \int_{x^4}^{x^2} f(x, y)\,dy$.

11. $\int_{-2}^{-1} dy \int_{-\sqrt{2+y}}^{0} f(x, y)\,dx + \int_{-1}^{0} dy \int_{-\sqrt{-y}}^{0} f(x, y)\,dx$.

12. $\int_0^{1/\sqrt{2}} dy \int_0^{\sin^{-1} y} f(x, y)\,dx + \int_{1/\sqrt{2}}^{1} dy \int_0^{\cos^{-1} y} f(x, y)\,dx$.

Draw the region of integration and take the following repeated integrals:

13. $\int_0^1 dx \int_0^x \sqrt{x+y}\,dy$. 14. $\int_{-1}^{1} dy \int_{y}^{y^2+y} xy\,dx$.

Calculate the area bounded by the lines:

15. $y^2 = 4ax$, $y = 0$, $x + y = 3a$. 16. $y = x$, $x^2 - 2ax = ay$. 17. $x = 4$, $y = x$, $xy = 4$. 18. $x - y = 1$, $y = -1$, $y = \ln x$. 19. $y = \sin x$, $y = \cos x$, $x = 0$. 20. $x = -6$, $x = -2$, $y = \frac{1}{x}$.

Calculate the area of one loop of the curves:

21. $\varrho = a \sin 2\varphi$. 22. $\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)^2 = \frac{2xy}{c^2}$.

Hint: Change the variables $x = a\varrho\cos\varphi$, $y = b\varrho\sin\varphi$.

Take the following integrals by passing over to polar coordinates.

23. $\iint_D (x^2+y^2)\,dx\,dy$ if D is bounded by the circle $x^2 + y^2 = 2ax$, $a > 0$.

24. $\iint\limits_D \frac{\ln(x^2 + y^2)}{x^2 + y^2}\,dx\,dy$ if D is the annulus between circles of radii $r = 1$ and $R = e$ with centre at the origin of coordinates.

25. $\int\limits_0^{2a} \int\limits_a^{\sqrt{2ax - x^2}} dy\,dx$. 26. $\int\limits_0^a dx \int\limits_0^{\sqrt{a^2 - x^2}} \sqrt{x^2 + y^2}\,dy$.

27. $\iint\limits_D y\,dx\,dy$, where D is a semicircle of diameter d with centre at $C(d/2, 0)$ lying above the x-axis.

Find the mass of a plate D bounded by the following curves and having the following surface density distribution $\gamma = \gamma(x, y)$:

28. D: $\{x = 1,\ y = 0,\ y^2 = x\ (y \geqslant 0)\}$, $\gamma = 3x - 6y^2$.

29. D: $\{x^2 + y^2 = 9,\quad x^2 + y^2 = 25,\quad x = 0,\quad y = 0\quad (x \leqslant 0,\quad y \leqslant 0)\}$, $\gamma = \frac{2y - x}{x^2 + y^2}$.

30. D: $\{x = 2,\ y = 0,\ y^2 = \frac{x}{2}\ (y \geqslant 0)\}$, $\gamma = 2x + 3y^2$.

Find the centre of mass of the figures:

31. Parabola's half-segment $y^2 = ax$, $x = a$, $y = 0$ $(y \geqslant 0)$.

32. Half-ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ cut by the x-axis.

33. Patch bounded by the lines $x + y = 2a$, $x^2 = ay$ $(a > 0)$.

Compute the areas of:

34. The part of the plane $x + y + z = 2a$ which lies in the first octant and is bounded by the cylinder $x^2 + y^2 = a^2$.

35. The part of the conic surface $x^2 + y^2 = z^2$ which is cut away by the cylinder $x^2 + y^2 = 2ax$.

36. The surface of the paraboloid $x^2 + y^2 = 2az$ that lies within the cylinder $x^2 + y^2 = 3a^2$.

Take the following surface integrals:

37. $\iint\limits_\pi xyz\,d\sigma$ where π is the part of the plane $x + y + z = 1$ lying in the first octant.

38. $\iint\limits_\pi x\,d\sigma$, where π is the part of the sphere $x^2 + y^2 + z^2 = R^2$ lying in the first octant.

39. $\iint\limits_\pi \frac{1}{r^2}\,d\sigma$, where π is the cylinder $x^2 + y^2 = R$ bounded by the planes $z = 0$ and $z = H$, and r is the distance from the surface π to the origin of coordinates.

Compute the moments of inertia of the following figures:

40. Triangle bounded by the lines $2y + x = 0$, $x = a$, $y = a$ relative to the x-axis.
41. Triangle with vertices at points $A(0, 2a)$, $B(a, 0)$, $C(a, a)$ relative to the y-axis.
42. Ellipse $x^2/a^2 + y^2/b^2 = 1$ relative to the y-axis.
43. **Definition.** The *moment of inertia* of a plane figure relative to the origin of coordinates is the quantity $I_0 = \iint\limits_D (x^2 + y^2)\,dx\,dy$. Find I_0 for the figure bounded by the parabola $y^2 = 4ax$, the straight line $y = 2a$ and the y-axis $(a > 0)$.

Compute the triple integrals:

44. $\iiint\limits_\Omega \frac{dx\,dy\,dz}{(x + y + z + 1)^3}$, where Ω is bounded by the coordinate planes and the plane $x + y + z = 1$.
45. $\iiint\limits_\Omega z\,dx\,dy\,dz$, where Ω is bounded by the cone $x^2 + y^2 = R^2z^2/H^2$ and the plane $z = H$.
46. $\iiint\limits_\Omega (2x + 3y - z)\,dx\,dy\,dz$, where Ω is a trigonal prism bounded by the planes $x = 0$, $y = 0$, $z = 0$, $z = a$, $x + y = b$ $(a > 0,\ b > 0)$.

Take the following integrals passing to cylindrical or spherical coordinates:

47. $\int\limits_0^1 dx \int\limits_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy \int\limits_0^a dz.$ 48. $\int\limits_0^2 dx \int\limits_0^{\sqrt{2x-x^2}} dy \int\limits_0^a z\sqrt{x^2 + y^2}\,dz.$

49. $\int\limits_0^1 dx \int\limits_0^{\sqrt{1-x^2}} dy \int\limits_0^{\sqrt{1-x^2-y^2}} \sqrt{x^2 + y^2 + z^2}\,dz.$

50. $\iiint\limits_\Omega \frac{dx\,dy\,dz}{\sqrt{x^2 + y^2 + (z - 2)^2}}$, where Ω is the sphere $x^2 + y^2 + z^2 \leqslant 1$.

Using triple integration compute the volumes of the bodies bounded by the following surfaces:

51. $x^2 + y^2 = z^2$, $x^2 + y^2 = z - 6$.
52. $az = x^2 + y^2$, $z^2 = x^2 + y^2$ $(a > 0)$.
53. $(x^2 + y^2 + z^2)^2 = axyz$ $(a > 0)$.

Hint: Pass over to spherical coordinates.

Compute the mass of the bodies:

54. Bounded by the surfaces $x^2 + y^2 = z^2$, $z = H$, if the density μ at each point of the body is equal to the z-coordinate of the point.
55. Bounded by the surfaces $x + z = a$, $y = 0$ $(y > 0)$, $y^2 = ax$, $2x + z = 2a$, if the density at each point is equal to the y-coordinate of the point.

Find the static moments of the following uniform ($\mu = 1$) bodies:

56. Rectangular parallelepiped with sides a, b, c relative to its faces.
57. A body bounded by the ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ and the xy-plane relative to the xy-plane.

Find the centres of mass of uniform ($\mu = 1$) bodies:

58. Bounded by the planes $x = 0$, $y = 0$, $z = 0$, $x = 2$, $y = 4$, $x + y + z = 8$.
59. Bounded by the cylinder $y^2 = 2z$ and the planes $x = 0$, $y = 0$, $z = 0$, $2x + 3y = 12$.
60. Bounded by the paraboloid $x^2 + y^2 = 2az$ and the sphere $x^2 + y^2 + z^2 = 3a^2$ $(z \geqslant 0)$.

Answers

1. 156. 2. $\ln\frac{25}{24}$. 3. $\frac{3\pi b^2}{16}$. 4. $\frac{2}{3}$. 5. $\int\limits_1^2 dx \int\limits_3^4 f(x, y)dy$. 6. $\int\limits_0^a dy \int\limits_0^{\sqrt{a^2-y^2}} f(x, y)dx$.
7. $\int\limits_0^1 dy \int\limits_{y/2}^{3y} f(x, y)dx + \int\limits_1^6 dy \int\limits_{y/2}^{3-y} f(x, y)dx$. 8. $\int\limits_{-1}^0 dx \int\limits_{-\sqrt{4x+4}}^{\sqrt{4x+4}} f(x, y)dy + \int\limits_0^8 dx \int\limits_{-\sqrt{4x+4}}^{2-x} f(x, y)dy$.
9. $\int\limits_{-4}^0 dx \int\limits_{-\sqrt{x+4}}^0 dy$. 10. $\int\limits_0^1 dy \int\limits_{\sqrt{y}}^{\sqrt[4]{y}} f(x, y)dx$. 11. $\int\limits_{-1}^0 dx \int\limits_{x^2-2}^{-x^2} f(x, y)dy$. 12. $\int\limits_0^{\frac{\pi}{4}} dx \int\limits_{\sin y}^{\cos y} f(x, y)dy$.
13. $4(2\sqrt{2} - 1)/15$. 14. $\frac{2}{5}$. 15. $\frac{10}{3}a^2$. 16. $\frac{9}{2}a^2$. 17. $6 - 4\ln 2$. 18. $\frac{e-2}{2e}$. 19. $\sqrt{2} - 1$. 20. $\ln 3$.
21. $\frac{\pi a^2}{8}$. 22. $\left(\frac{ab}{c}\right)^2$. 23. $\frac{3\pi}{2}a^4$. 24. 2π. 25. $\frac{\pi a^2}{2}$. 26. $\frac{\pi a^3}{6}$. 27. $\frac{d^3}{12}$. 28. 2. 29. 6. 30. 4.
31. $x_c = \frac{3a}{5}$, $y_c = \frac{3a}{8}$. 32. $x_c = 0$, $y_c = \frac{4c}{3\pi}$. 33. $x_c = -\frac{a}{2}$, $y_c = \frac{8a}{5}$. 34. $\frac{\sqrt{3}}{4}\pi a^2$.
35. $2\sqrt{2}\pi a^2$. 36. $\frac{14}{3}\pi a^2$. 37. $\frac{\sqrt{3}}{120}$. 38. $\frac{\pi R^3}{4}$. 39. $2\pi\tan^{-1}\frac{H}{R}$. 40. $\frac{17a^4}{96}$. 41. $\frac{a^4}{4}$. 42. $\frac{\pi a^3 b}{4}$.
43. $\frac{178}{165}a^4$. 44. $\frac{1}{2}\left(\ln 2 - \frac{5}{8}\right)$. 45. $\frac{\pi H^2 R^2}{4}$. 46. $\frac{5}{6}ab^3 - \frac{1}{4}a^2b^2$. 47. $\frac{\pi a}{2}$. 48. $\frac{8}{9}a^2$. 49. $\frac{\pi}{8}$.
50. $\frac{2\pi}{3}$. 51. $\frac{32}{3}\pi$. 52. $\frac{\pi a^3}{6}$. 53. $\frac{a^3}{360}$. 54. $\frac{\pi H^4}{4}$. 55. $\frac{a^4}{12}$. 56. $\frac{a^2bc}{2}$, $\frac{ab^2c}{2}$, $\frac{abc^2}{2}$.
57. $\frac{\pi abc^2}{4}$. 58. $x_c = \frac{14}{15}$, $y_c = \frac{26}{15}$, $z_c = \frac{8}{3}$. 59. $x_c = \frac{6}{5}$, $y_c = \frac{12}{5}$, $z_c = \frac{8}{5}$. 60. $x_c = 0$, $y_c = 0$, $z_c = \frac{5a}{83}(6\sqrt{3} + 5)$.

Chapter 23

Line Integrals

23.1 Line Integrals of the First Kind

A curve AB given by the parametric equations

$$\begin{cases} x = \varphi(t) \\ y = \psi(t) \end{cases} \qquad (t_0 \leqslant t \leqslant t_1)$$

is called *smooth*, if the functions $\varphi(t)$ and $\psi(t)$ have continuous derivatives $\varphi'(t)$ and $\psi'(t)$ on the interval $[t_0, t_1]$.

A continuous curve composed of a finite number of smooth pieces is said to be *piecewise smooth.*

Let AB be a plane curve, smooth or piecewise smooth, and let $f(M)$ be a function defined on AB or a domain D that contains AB. Consider

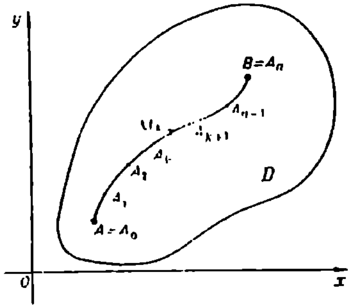


Fig. 23.1

some dissection of the curve AB by the points $A = A_0, A_1, A_2, \ldots, A_n = B$ (Fig. 23.1). On each arc A_kA_{k+1} we take an arbitrary point M_k and set up the sum

$$\sigma = \sum_{k=0}^{n-1} f(M_k)\Delta l_k, \tag{23.1}$$

where Δl_k is the length of the arc A_kA_{k+1}.

The sum (23.1) is known as the *integral sum for $f(M)$ along the arc of the curve.* Let Δl be the largest of the lengths of those partial arcs, i.e.,

$$\Delta l = \max_{0 \leqslant k \leqslant n-1} \Delta l_k .$$

Definition. If (23.1) has a finite limit when $\Delta l \to 0$, which is dependent neither on the way in which AB is broken into parts nor points M_k are chosen, then this limit is called the *line integral of the first kind of $f(M)$ along the curve AB* (*the integral along the arc of the curve*) and is denoted by

$$\int_{AB} f(M)\, dl \qquad \text{or} \qquad \int_{AB} f(x, y)\, dl$$

(the point $(x, y) \in AB$). Thus, by definition,

$$\int_{AB} f(M)\, dl = \lim_{\Delta l \to 0} \sum_{k=0}^{n-1} f(M_k)\, \Delta l_k . \tag{23.2}$$

The function $f(M)$ is then called *integrable along the curve AB*, and AB is called the *contour of integration, A the initial point* and B the *final point of integration.*

Example. Along a smooth curve L a mass m is distributed with a linear density $f(M)$. Find m.

◄ We break up L into n arbitrary parts $M_k M_{k+1}$ ($k = 0, 1, \ldots, n-1$) and estimate the mass of each part $M_k M_{k+1}$ assuming that in each of them the density is constant and is equal to the density at some point within them, e.g., at the leftmost point $f(M_k)$. Then,

$$\sum_{k=0}^{n-1} f(M_k)\, \Delta l_k ,$$

where Δl_k is the length of the kth part, will be an approximate value of the mass m. Clearly, the error will be the smaller the more detailed the dissection of L. In the limit as $\Delta l \to 0$ ($\Delta l = \max_k \Delta l_k$) we will obtain the exact value of the mass of L, i.e.,

$$m = \lim_{\Delta l \to 0} \sum_{k=0}^{n-1} f(M_k)\, \Delta l_k .$$

But the limit on the right is a line integral of the first kind. Hence,

$$m = \int_{AB} f(M)\, dl. \ ►$$

Now we prove the existence of the line integral of the first kind. For curve AB we will regard as a parameter the length l of the arc reckoned

from the initial point A (Fig. 23.2). The curve AB will then be given by

$$\begin{cases} x = x(l) \\ y = y(l) \end{cases} \quad 0 \leqslant l \leqslant L, \tag{23.3}$$

where L is the length of AB.

Equations (23.3) are called the *natural equations* of AB. The function defined on the curve AB will then become a function of l: $f(x(l), y(l))$.

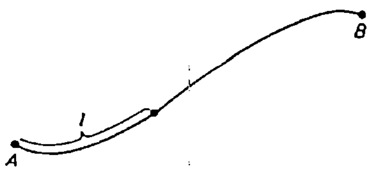


Fig. 23.2

We denote by l_k ($k = 0, 1, ..., n - 1$) the value of l corresponding to the point M_k, and rewrite the integral sum (23.1) as

$$\sum_{k=0}^{n-1} f(x(l_k), y(l_k)) \Delta l_k. \tag{23.4}$$

This is the integral sum that corresponds to the definite integral

$$\int_0^L f(x(l), y(l)) dl.$$

Since the integral sums (23.1) and (23.4) are equal, so are the integrals corresponding to them. Thus,

$$\int_{AB} f(M)\, dl = \int_0^L f(x(l), y(l))\, dl. \tag{23.5}$$

Theorem 23.1. *If a function $f(M)$ is continuous along a smooth curve AB then the line integral $\int_{AB} f(M)\, dl$ exists* (since under these conditions exists the definite integral on the right-hand side of (23.5)).

Properties of line integrals of the first kind. (1) It follows from the integral sum (23.1) that

$$\int_{AB} f(M)\, dl = \int_{BA} f(M)\, dl,$$

i.e., the value of a line integral of the first kind is independent of the direction of integration.

(2) *Linearity.* If for either of the functions $f(M)$ and $g(M)$ there exists a line integral over a curve AB and if α and β are arbitrary constants, then for the function $\alpha f(M) + \beta g(M)$ there exists a line integral along AB, such that

$$\int_{AB} [\alpha f(M) + \beta g(M)]dl = \alpha \int_{AB} f(M)\,dl + \beta \int_{AB} g(M)\,dl.$$

(3) *Additivity.* If AB is made up of two pieces AC and CB and for $f(M)$ there exists a line integral along AB, then there exist the integrals $\int_{AC} f(M)\,dl$ and $\int_{CB} f(M)\,dl$, such that

$$\int_{AB} f(M)\,dl = \int_{AC} f(M)\,dl + \int_{CB} f(M)\,dl.$$

(4) If $f(M) \geqslant 0$ on the curve AB, then

$$\int_{AB} f(M)\,dl \geqslant 0.$$

(5) If $f(M)$ is integrable along AB, then the function $|f(M)|$ is also integrable along AB, and

$$\left|\int_{AB} f(M)\,dl\right| \leqslant \int_{AB} |f(M)|\,dl.$$

(6) *Mean value formula.* If a function $f(M)$ is continuous along a curve AB, then on this curve there is a point M_m such that

$$\int_{AB} f(M)\,dl = f(M_m)\,L.$$

where L is the length of AB.

Computation of the line integral of the first kind. Let a curve AB be given by the following parametric equations:

$$\begin{cases} x = \varphi(t) \\ y = \psi(t) \end{cases} \quad t_0 \leqslant t \leqslant t_1.$$

Corresponding to the point A is the value $t = t_0$, and to the point B the value $t = t_1$. We will assume that the functions $\varphi(t)$ and $\psi(t)$ are continuous on $[t_0, t_1]$ together with their derivatives $\varphi'(t)$ and $\psi'(t)$, and

$$[\varphi'(t)]^2 + [\psi'(t)]^2 > 0,$$

Then the differential of the arc of the curve will be

$$dl = \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}\,dt$$

and

$$\int\limits_{AB} f(x, y)\,dl = \int\limits_{t_0}^{t_1} f(\varphi(t), \psi(t))\sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2}\,dt.$$

Specifically, if AB is given by the explicit equation $y = g(x)$, $a \leqslant x \leqslant b$, where $g(x)$ is continuously differentiable on $[a, b]$ and corresponding to the point A is the value $x = a$, and to the point B is the value $x = b$, then if we assume x to be the parameter, we will arrive at

$$\int\limits_{AB} f(x, y)\,dl = \int\limits_{a}^{b} f(x, g(x))\sqrt{1 + [g'(x)]^2}\,dx.$$

Line integrals of the first kind in space. The definition of the line integral of the first kind formulated above for a plane curve can literally be applied to the case of a function $f(M)$ defined along a space curve AB.

If the curve is given by the parametric equations

$$\begin{cases} x = \varphi(t) \\ y = \psi(t) \qquad t_0 \leqslant t \leqslant t_1. \\ z = \omega(t) \end{cases}$$

then the line integral of the first kind along this curve is reduced to a definite integral by

$$\int\limits_{AB} f(x, y, z)\,dl$$

$$= \int\limits_{t_0}^{t_1} f[\varphi(t), \psi(t), \omega(t)]\sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2 + [\omega'(t)]^2}\,dt.$$

Example. Take the line integral $\int\limits_{L} (x + y)\,dl$, where L is the contour of a triangle with vertices at the points $O(0, 0)$, $A(1, 0)$, and $B(0, 1)$ (Fig. 23.3).

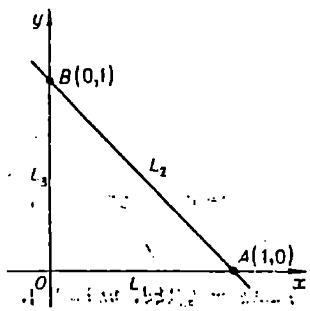


Fig. 23.3

◄ By the additivity property we have

$$\int_L (x+y)\,dl = \int_{OA} (x+y)\,dl + \int_{AB} (x+y)\,dl + \int_{BO} (x+y)\,dl.$$

We take each integral separately. Since on OA we have $0 \leqslant x \leqslant 1$, $y = 0$ and $dl = dx$, then

$$\int_{OA} (x+y)\,dl = \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}.$$

On AB we have $x + y = 1$, or $y = 1 - x$. We thus obtain $y' = -1$ and $dl = \sqrt{1 + (y')^2}\,dx = \sqrt{2}\,dx$, $0 \leqslant x \leqslant 1$, then

$$\int_{AB} (x+y)\,dl = \int_0^1 \sqrt{2}\,dx = \sqrt{2}.$$

Lastly,

$$\int_{BO} (x+y)\,dl = \int_{OB} (x+y)\,dl = \int_0^1 y\,dy = \frac{y^2}{2}\Big|_0^1 = \frac{1}{2}.$$

Consequently,

$$\int_L (x+y)\,dl = \frac{1}{2} + \sqrt{2} + \frac{1}{2} = 1 + \sqrt{2}.$$

Remark. Computing the integrals $\int_{AB} (x+y)\,dl$ and $\int_{BO} (x+y)\,dl$ we made use of property (1), according to which

$$\int_{AB} (x+y)dl = \int_{BA} (x+y)dl \quad \text{and} \quad \int_{BO} (x+y)dl = \int_{OB} (x+y)dl.$$

23.2 Line Integrals of the Second Kind

Let AB be a smooth or piecewise smooth oriented curve in the xy-plane and let $\mathbf{F}(M) = P(M)\mathbf{i} + Q(M)\mathbf{j}$ be a vector function defined in a certain domain D that contains the curve AB.

We will break up AB into parts by points $A = A_0, A_1, \ldots, A_n = B$, whose coordinates we will denote by $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, respectively. We will take on each elementary arc A_kA_{k+1} an arbitrary point $M_k(\xi_k, \eta_k)$ and form the sum

$$\sigma = \sum_{k=0}^{n-1} [P(\xi_k, \eta_k)\Delta x_k + Q(\xi_k, \eta_k)\Delta y_k], \tag{23.6}$$

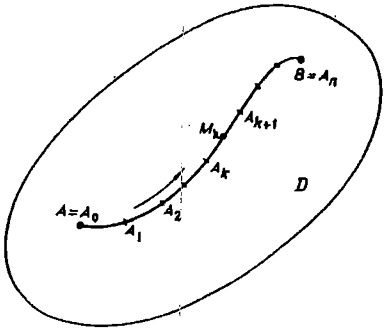


Fig. 23.4

where $\Delta x_k = x_{k+1} - x_k$, $\Delta y_k = y_{k+1} - y_k$ $(k = 0, 1, 2, \ldots, n-1)$. Let Δl be the largest of the arcs A_kA_{k+1} (Fig. 23.4).

Definition. If the sum (23.6) has a finite limit as $\Delta l \to 0$, which is dependent neither on the way in which AB has been divided nor on how points (ξ_k, η_k) have been selected on elementary arcs, then this limit is called the *line integral of the second kind* of the vector function $\mathbf{F}(M)$ along the curve AB. We will denote it by

$$\int_{AB} P(x, y)\,dx + Q(x, y)\,dy.$$

And so, by definition,

$$\int_{AB} P(x, y)\,dx + Q(x, y)\,dy$$
$$= \lim_{\Delta l \to 0} \sum_{k=0}^{n-1} [P(\xi_k, \eta_k)\Delta x_k + Q(\xi_k, \eta_k)\Delta y_k]. \tag{23.7}$$

Theorem 23.2. *If in a certain domain D that contains a curve AB the functions $P(x, y)$ and $Q(x, y)$ are continuous, then there exists the line integral of the second kind* $\int_{AB} P(x, y)\,dx + Q(x, y)\,dy$.

Let a vector $\mathbf{r}(M) = x\mathbf{i} + y\mathbf{j}$ be the radius vector of a point $M(x, y)$. Then $d\mathbf{r} = \mathbf{i}dx + \mathbf{j}dy$, and the integrand $P(x, y)\,dx + Q(x, y)\,dy$ in (23.7) can be represented as the scalar product of $\mathbf{F}(M)$ and $d\mathbf{r}$. Therefore, the integral of the second kind of $\mathbf{F}(M) = P(M)\mathbf{i} + Q(M)\mathbf{j}$ along AB can for short be written as $\int_{AB} (\mathbf{F}, d\mathbf{r})$.

Computation of the line integral of the second kind. Suppose that a curve AB is given by the parametric equations

$$\begin{cases} x = \varphi(t) \\ y = \psi(t) \end{cases} \quad t_0 \leqslant t \leqslant t_1;$$

where the functions $\varphi(t)$ and $\psi(t)$ are continuous together with the derivatives $\varphi'(t)$, $\psi'(t)$ on an interval $[t_0, t_1]$, and as t varies from t_0 to t_1 the point $M(x, y)$ moves along AB from A to B.

If in a certain domain D that contains the curve AB the functions $P(x, y)$ and $Q(x, y)$ are continuous, then the line integral of the second kind

$$\int_{AB} P(x, y)\,dx + Q(x, y)\,dy$$

reduces to the following definite integral:

$$\int_{AB} P(x, y)\,dx + Q(x, y)\,dy = \int_{t_0}^{t_1} [P(\varphi(t), \psi(t))\varphi'(t) + Q(\varphi(t), \psi(t))\,\psi'(t)]\,dt. \tag{23.8}$$

To sum up, to take a line integral of the second kind means to take a conventional definite integral.

Example. Take the integral $\int_{AB} x\,dy - y\,dx$ (1) along the straight segment connecting the points $A(0, 0)$ and $B(1, 1)$; (2) along the parabola $y = x^2$ connecting the same points (Fig. 23.5).

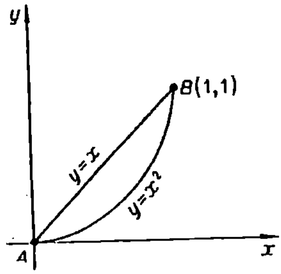


Fig. 23.5

◄ (1) The equation of AB is $y = x$ (x is the parameter, $0 \leqslant x \leqslant 1$), hence $dy = dx$. Thus,

$$\int_{AB} x\,dy - y\,dx = \int_0^1 (x\,dx - x\,dx) = 0.$$

(2) The equation of AB is $y = x^2$, $0 \leqslant x \leqslant 1$, hence $dy = 2x\,dx$. Therefore, $x\,dy = 2x^2\,dx$ and

$$\int_{AB} x\,dy - y\,dx = \int_0^1 2x^2\,dx - x^2\,dx = \int_0^1 x^2\,dx = \frac{1}{3}. \;►$$

Thus in the general case the value of a line integral of the second kind depends on the shape of integration path.

Properties of line integrals of the second kind. (1) *Linearity.* If there exist line integrals

$$\int_{AB} (\mathbf{F}_1, d\mathbf{r}) \quad \text{and} \quad \int_{AB} (\mathbf{F}_2, d\mathbf{r}),$$

then for any real α and b there exists the integral

$$\int_{AB} (\alpha \mathbf{F}_1 + \beta \mathbf{F}_2, d\mathbf{r}),$$

such that

$$\int_{AB} (\alpha \mathbf{F}_1 + \beta \mathbf{F}_2, d\mathbf{r}) = \alpha \int_{AB} (\mathbf{F}_1, d\mathbf{r}) + \beta \int_{AB} (\mathbf{F}_2, d\mathbf{r}).$$

(2) *Additivity.* If a curve AB is made up of parts AC and CB and there exists the line integral $\int_{AB} (\mathbf{F}, d\mathbf{r})$, then there exist the integrals $\int_{AC} (\mathbf{F}, d\mathbf{r})$ and $\int_{CB} (\mathbf{F}, d\mathbf{r})$, such that

$$\int_{AB} (\mathbf{F}, d\mathbf{r}) = \int_{AC} (\mathbf{F}, d\mathbf{r}) + \int_{CB} (\mathbf{F}, d\mathbf{r}).$$

(3) Unlike the line integral of the first kind, the integral of the second kind depends on the direction (from A to B or from B to A) in which curve AB is passed, and it changes its sign when the direction is changed, i.e.,

$$\int_{BA} P\,dx + Q\,dy = -\int_{AB} P\,dx + Q\,dy.$$

◄ If we change the direction of the passage of AB, we thereby will change Δx_k and Δy_k in the integral sum

$$\sum_{k=0}^{n-1} P\,\Delta x_k + Q\,\Delta y_k$$

for $(-\Delta x_k)$ and $(-\Delta y_k)$, respectively. This will change the sign of the sum, and hence of its limit. ►

This property corresponds to the physical interpretation of the line integral of the second kind as the work done by a field of forces $\mathbf{F}$ along some path: changing the direction of motion along a curve changes the sign of the work done by the field along the curve.

Relationship between the line integrals of the first and second kinds. Consider the line integral of the second kind

$$\int (\mathbf{F}, d\mathbf{r})$$

where AB is an oriented curve (A is the initial point and B is the finite point) given by the following vector-parametric equation

$$\mathbf{r} = \mathbf{r}(l),$$

where l is the length of the curve reckoned in the direction of orientation of AB (Fig. 23.6). Then

$$\frac{d\mathbf{r}}{dl} = \tau \qquad \text{or} \qquad d\mathbf{r} = \tau\, dl,$$

where τ is the unit vector of the tangent to AB at a point M. Therefore,

$$\int_{AB} (\mathbf{F}, d\mathbf{r}) = \int_{AB} (\mathbf{F}, \tau\, dl) = \int_{AB} (\mathbf{F}, \tau)\, dl,$$

where on the right we have a line integral along an arc of the curve, i.e., a line integral of the first kind.

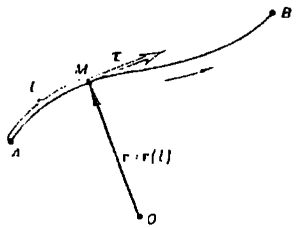


Fig. 23.6

When the orientation of the curve AB changes, the unit vector of the tangent τ is replaced by the opposite unit vector $(-\tau)$, which in turn changes the sign of the integral.

23.3 Green's Formula

We introduce Green's formula relating the line integral

$$\int_L P(x, y)\, dx + Q(x, y)\, dy$$

along the boundary L of a domain D to a double integral over the same domain.

Theorem 23.3. *If in a closed domain D bounded by a piecewise smooth contour L the functions $P(x, y)$ and $Q(x, y)$ are continuous and have continuous partial derivatives $\frac{\partial Q}{\partial x}$ and $\frac{\partial P}{\partial y}$, then we have the following rela-*

tions (Green's formula)

$$\oint_L P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy. \tag{23.9}$$

Here $\oint_L$ stands for integration along the boundary L of the domain D, which is passed so that D lies on the left (Fig. 23.7).

Remark. If L consists of a finite number of piecewise smooth curves L_i, then the domain is called *multiply connected,* and L_i are called connected components of the boundary. Figure 23.8 shows a triply connected region.

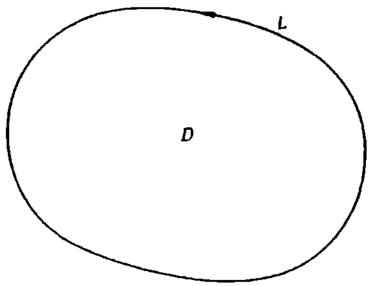


Fig. 23.7

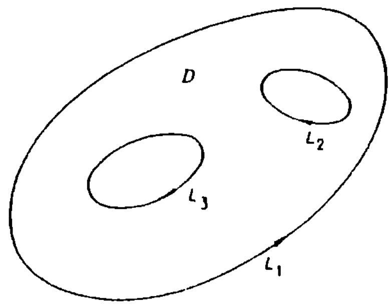


Fig. 23.8

We will say that a domain D is simply connected if it includes no "holes", i.e., it is a region such that any closed curve drawn in it can be contracted into a point $P \in D$, while remaining in D.

◄ We will prove this for a simply connected region. By the linearity property it is sufficient to prove that

$$\oint_L P(x, y)\,dx = -\iint_D \frac{\partial P}{\partial y}\,dx\,dy, \tag{23.10}$$

$$\oint_L Q(x, y)\,dy = \iint_D \frac{\partial Q}{\partial x}\,dx\,dy. \tag{23.11}$$

Let us prove the first of these. Suppose at first that L is cut across by each straight line parallel to the y-axis at no more than two points (Fig. 23.9). This divides L into two parts (upper and lower), each of which is projected in a one-to-one manner on an interval $[a, b]$ on the x-axis.

be approximately $f(M_k)\,\Delta l_k$, and the area of the surface $ABDC$ will be

$$S = \sum_{k=0}^{n-1} f(M_k)\,\Delta l_k.$$

The approximate equality will be the more exact the smaller will be the partial arcs M_kM_{k+1} into which the curve AB is divided. Let Δl be the largest of the lengths Δl_k of M_kM_{k+1}. In the limit as $\Delta l \to 0$ we then obtain the exact value of the area

$$S = \lim_{\Delta l\to 0} \sum_{k=0}^{n-1} f(M_k)\,\Delta l_k.$$

By definition, the limit on the right-hand side is the line integral of the first kind of $f(M)$ along the curve AB. We thus arrive at

$$S = \int_{AB} f(M)\,dl. \tag{23.17}$$

Example. Calculate the part of the cylindrical surface $x^2 + y^2 = R^2$ ($x \geqslant 0$, $y \geqslant 0$) cut from above by the surface $xy = 2Rz$.

◄ We reduce the problem to taking the line integral of the first kind of the function $z = xy/2R$ along the arc of the circle lying in the first quadrant. We will have

$$S = \int_{AB} \frac{xy}{2R}\,dl.$$

The parametric equations of AB are

$$x = R\cos t, \qquad y = R\sin t, \qquad 0 \leqslant t \leqslant \frac{\pi}{2}.$$

Then

$$dl = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt = \sqrt{R^2\sin^2 t + R^2\cos^2 t}\,dt = R\,dt$$

and

$$S = \int_0^{\pi/2} \frac{R\cos t\cdot R\sin t}{2R}\,R\,dt = \frac{R^2}{2}\,\frac{\sin^2 t}{2}\bigg|_0^{\pi/2} = \frac{R^2}{4}. \text{ ►}$$

Area of a plane figure. When we discussed the special form of Green's formula, we established that the area S of a plane figure D bounded by a line L is given by

$$S = \frac{1}{2}\oint_L x\,dy - y\,dx. \tag{23.18}$$

The right-hand side is a line integral of the second kind.

Work done by a force. Suppose that in a plane domain D containing a line AD we specify a force

$$\mathbf{F}(M) = P(M)\mathbf{i} + Q(M)\mathbf{j}, \tag{23.19}$$

where $P(M)$ and $Q(M)$, and hence $\mathbf{F}(M)$, are assumed to be continuous functions of the point M. We would like to find the work done by the force $\mathbf{F}$, if under the action of it the material point M of a unit mass moved from point A to point B along the curve AB.

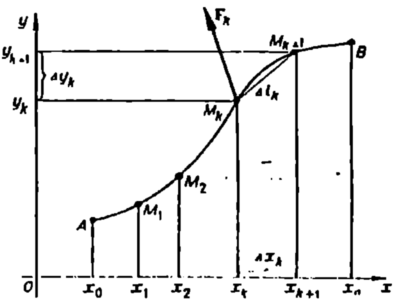


Fig. 23.12

We break up AB in an arbitrary way into n parts by points $A = M_0$, $M_1, \ldots, M_k, M_{k+1}, \ldots, M_n = B$ (Fig. 23.12). We then replace each arc $\overset{\frown}{M_kM_{k+1}}$ by the chord $\overline{M_kM_{k+1}}$ and suppose for simplicity that on $\overset{\frown}{M_kM_{k+1}}$ (and hence on $\overline{M_kM_{k+1}}$) the force $\mathbf{F}_k$ has a constant value, e.g., the value at M_k:

$$\mathbf{F}_k = \mathbf{F}(M_k). \tag{23.20}$$

We thus obtain the approximate expression for the work done by the force on $\overset{\frown}{M_kM_{k+1}}$

$$W_k = |\mathbf{F}_k| \cdot |\Delta \mathbf{l}_k| \cdot \cos(\widehat{\mathbf{F}_k, \Delta \mathbf{l}_k}), \tag{23.21}$$

where $|\mathbf{F}_k|$ is the length of the vector $\mathbf{F}_k$, and

$$\Delta \mathbf{l}_k = \overline{M_kM_{k+1}} = \Delta x_k \mathbf{i} + \Delta y_k \mathbf{j}. \tag{23.22}$$

From (23.19) we will obtain using (23.20)

$$\mathbf{F}_k = P(M_k)\mathbf{i} + Q(M_k)\mathbf{j}$$

or

$$\mathbf{F}_k = P(x_k, y_k)\mathbf{i} + Q(x_k, y_k)\mathbf{j}. \tag{23.23}$$

The right-hand side of (23.21) is a scalar product of $\mathbf{F}_k$ and $\Delta \mathbf{l}_k$, and so

we will have, by (23.22) and (23.23),

$$W_k \simeq (\mathbf{F}_k,\ \Delta \mathbf{l}_k) = P(x_k, y_k)\Delta x_k + Q(x_k, y_k)\Delta y_k.$$

Summing up over all the values of k ($k = 0, 1, 2, ..., n - 1$), we will get the quantity

$$W \simeq \sum_{k=0}^{n-1} P(x_k, y_k)\Delta x_k + Q(x_k, y_k)\,\Delta y_k,$$

which expresses approximately the work done by $\mathbf{F}(M)$ along the entire path from A to B. The limit of this sum as $\Delta x_k \to 0$ and $\Delta y_k \to 0$ is assumed to be the exact value of the work. On the other hand, however, the limit of the sum is a line integral of the second kind of the vector function $\mathbf{F}(M)$ along AB. The work done by the force is thus given by

$$W = \int_{AB} P(x, y)dx + Q(x, y)\,dy. \qquad (23.24)$$

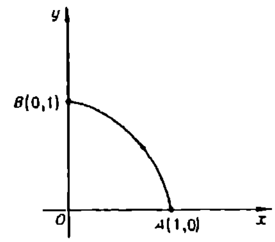


Fig. 23.13

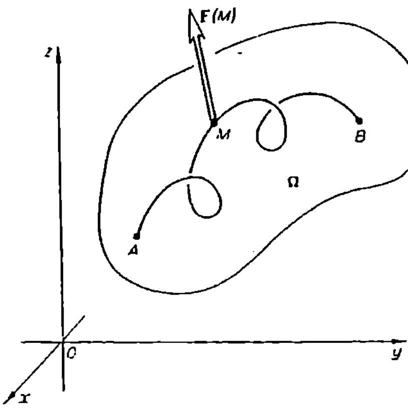


Fig. 23.14

Example. Given a force $\mathbf{F} = x^2\mathbf{j}$, find the work done by it in moving a unit mass along the parabola $y^2 = 1 - x$ from the point $A(1, 0)$ to the point $B(0, 1)$ (Fig. 23.13).

◄ We use formula (23.24), where we put $P(x, y) = 0$, $Q(x, y) = x^2$. The curve AB in this case is the parabola $y^2 = 1 - x$, so that $x = 1 - y^2$ and

$$W = \int_{AB} x^2 dy = \int_0^1 (1 - y^2)^2 dy = \int_0^1 (1 - 2y^2 + y^4)dy = \frac{8}{15}. \ ►$$

We can easily generalize the results to the case of the space curve (Fig. 23.14). In a space domain Ω that contains a space curve AB we specify

a force

$$\mathbf{F}(M) = P(M)\mathbf{i} + Q(M)\mathbf{j} + R(M)\mathbf{k},$$

where $P(M)$, $Q(M)$, and $R(M)$ are continuous functions in Ω, then the work done by $\mathbf{F}(M)$ to move a material point M of a unit mass from the point A to the point B along AB will be

$$W = \int_{AB} P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz.$$

Exercises

Take the following integrals of the first kind:

1. $\int_L xy\,dl$, where L is a quarter of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ lying in the first quadrant.
2. $\int_L (x - y)dl$, where L is the circle $x^2 + y^2 = 2ax$.
3. $\int_L \frac{dl}{x - y}$, where L is the segment connecting the points $(0, -2)$ and $(4, 0)$.
4. $\int_L x\,dl$, where L is the segment connecting the points $(0, 0)$ and $(1, 2)$.
5. $\int_L y\,dl$, where L is the arc of the parabola $y^2 = 2x$ from the point $(0, 0)$ to the point $(1, \sqrt{2})$.
6. $\int_L \frac{dl}{x^2 + y^2 + z^2}$, where L is the first turn of the helical line $x = a\cos t$, $y = a\sin t$, $z = bt$.
7. Find the length of the arc of the conical spiral $x = ae^t\cos t$, $y = ae^t\sin t$, $z = ae^t$ from the point $A(0, 0, 0)$ to the point $B(a, 0, a)$. *Hint:* Corresponding to A is $t_1 = -\infty$, and to B is $t_2 = 0$.
8. Find the lateral area of the circular cylindrical surface lying above the first turn of the spiral $x = a\cos t$, $y = a\sin t$, $z = bt$ and above the plane $z = 0$.
9. Find the coordinates of the centre of mass of the uniform semiarc of the cycloid $x = a(t - \sin t)$, $y = a(1 - \cos t)$, $(0 \leqslant t \leqslant \pi)$.

Take the following line integrals of the second kind:

10. $\int_L y\,dx + x\,dy$, where L is the arc of the curve $y = x^3$ from the point $(0, 0)$ to the point $(2, 8)$.
11. $\int_L y^2\,dx + x^2\,dy$, where L is the upper half of the ellipse $x = a\cos t$, $y = b\sin t$ passed counterclockwise.

This definition of the *directional derivative* is invariant, i.e., it is in no way associated with the choice of a coordinate system.

We find the expression for the directional derivative in a rectangular system of coordinates. Let a function $f(M) = f(x, y, z)$ be differentiable at the point $M_0(x_0, y_0, z_0)$. Consider the value of $f(M)$ at a point $M(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$. The total increment of the function can then be written as

$$\Delta u = f(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) - f(x_0, y_0, z_0)$$
$$= \left.\frac{\partial u}{\partial x}\right|_{M_0} \Delta x + \left.\frac{\partial u}{\partial y}\right|_{M_0} \Delta y + \left.\frac{\partial u}{\partial z}\right|_{M_0} \Delta z + \varepsilon \cdot \Delta l,$$

where $\varepsilon \to 0$ as $\Delta l = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \to 0$, and the symbols $\left.\frac{\partial u}{\partial x}\right|_{M_0}$, $\left.\frac{\partial u}{\partial y}\right|_{M_0}$, $\left.\frac{\partial u}{\partial z}\right|_{M_0}$ mean that the partial derivatives are taken at the point M_0.

Then

$$\left.\frac{\partial u}{\partial l}\right|_{M_0} = \left.\frac{\partial u}{\partial x}\right|_{M_0} \cdot \lim_{\Delta l \to 0} \frac{\Delta x}{\Delta l}$$
$$+ \left.\frac{\partial u}{\partial y}\right|_{M_0} \cdot \lim_{\Delta l \to 0} \frac{\Delta y}{\Delta l} + \left.\frac{\partial u}{\partial z}\right|_{M_0} \cdot \lim_{\Delta l \to 0} \frac{\Delta z}{\Delta l}. \tag{24.7}$$

Here the quantities $\Delta x/\Delta l$, $\Delta y/\Delta l$, and $\Delta z/\Delta l$ are the direction cosines of $\overrightarrow{M_0M} = \Delta x\mathbf{i} + \Delta y\mathbf{j} + \Delta z\mathbf{k}$. Since $\overrightarrow{M_0M} \uparrow\uparrow \mathbf{l}$, the vectors have the same direction cosines

$$\frac{\Delta x}{\Delta l} = \cos\alpha, \quad \frac{\Delta y}{\Delta l} = \cos\beta, \quad \frac{\Delta z}{\Delta l} = \cos\gamma,$$

where

$$\mathbf{l}^0 = \frac{\mathbf{l}}{|\mathbf{l}|} = \mathbf{i}\cos\alpha + \mathbf{j}\cos\beta + \mathbf{k}\cos\gamma.$$

Since $M \to M_0$ while it remains on a straight line parallel to $\mathbf{l}$, the angles α, β and γ are constant, and therefore,

$$\lim_{\Delta l \to 0} \frac{\Delta x}{\Delta l} = \cos\alpha, \quad \lim_{\Delta l \to 0} \frac{\Delta y}{\Delta l} = \cos\beta, \quad \lim_{\Delta l \to 0} \frac{\Delta z}{\Delta l} = \cos\gamma. \tag{24.8}$$

From (24.7) and (24.8) we arrive at

$$\left.\frac{\partial u}{\partial l}\right|_{M_0} = \left.\frac{\partial u}{\partial x}\right|_{M_0} \cos\alpha + \left.\frac{\partial u}{\partial y}\right|_{M_0} \cos\beta + \left.\frac{\partial u}{\partial z}\right|_{M_0} \cos\gamma. \tag{24.9}$$

Remark. The partial derivatives $\partial u/\partial x$, $\partial u/\partial y$ and $\partial u/\partial z$ are the derivatives of u with respect to the directions of the x-, y- and z-axes, respectively.

Examples. (1) Find the derivative of $u = xe^y + ye^x - z^2$ at the point $M_0(3, 0, 2)$ directed to a point $M_1(4, 1, 3)$.

◄ We have

$$\left.\frac{\partial u}{\partial x}\right|_{M_0} = \left.(e^y + ye^x)\right|_{M_0} = 1,$$

$$\left.\frac{\partial u}{\partial y}\right|_{M_0} = \left.(xe^y + e^x)\right|_{M_0} = 3 + e^3,$$

$$\left.\frac{\partial u}{\partial z}\right|_{M_0} = \left.-2z\right|_{M_0} = -4.$$

$\vec{M_0M_1} = \Delta x \vec{i} + \Delta y \vec{j} + \Delta$
$|M_0M_1| = (1, 1$
$\frac{\Delta x}{\Delta l} = \cos\alpha$
$\Delta l = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} =$
$\frac{1}{\sqrt{3}} = \cos\alpha$

The vector $\vec{M_0M} = \{1, 1, 1\}$ has the magnitude $|\vec{M_0M}| = \sqrt{3}$. Its direction cosines are $\cos\alpha = 1/\sqrt{3}$, $\cos\beta = 1/\sqrt{3}$, $\cos\gamma = 1/\sqrt{3}$. By (24.9) we will have

$$\left.\frac{\partial u}{\partial l}\right|_{M_0} = 1\,\frac{1}{\sqrt{3}} + (3 + e^3)\,\frac{1}{\sqrt{3}} - 4\,\frac{1}{\sqrt{3}} = \frac{e^3}{\sqrt{3}}.$$

The fact that $\left.\frac{\partial u}{\partial l}\right|_{M_0} > 0$ means that the scalar field at M_0 grows in the given direction.

For a plane field $u = f(x, y)$ the derivative in the direction $\mathbf{l}$ at the point $M_0(x_0, y_0)$ is computed by

$$\left.\frac{\partial u}{\partial l}\right|_{M_0} = \left.\frac{\partial u}{\partial x}\right|_{M_0} \cos\alpha + \left.\frac{\partial u}{\partial y}\right|_{M_0} \sin\alpha, \tag{24.10}$$

where α is the angle between $\mathbf{l}$ and the x-axis. ►

(2) Calculate the derivative of the scalar field $u = \tan^{-1} xy$ at the point $M_0(1, 1)$ lying on the parabola $y = x^2$ in the direction along the curve (in the direction of increasing the abscissa).

◄ The direction $\mathbf{l}$ of the parabola $y = x^2$ at $M_0(1, 1)$ is the direction of the tangent to the parabola at that point (Fig. 24.3).

Suppose that the tangent to the parabola at M_0 forms with the x-axis an angle α. Then $\tan\alpha = y'(x)\big|_{x=1} = 2$, whence the direction cosines of the tangents will be

$$\cos\alpha = \frac{1}{\sqrt{1 + \tan^2\alpha}} = \frac{1}{\sqrt{5}}, \quad \sin\alpha = \frac{\tan\alpha}{\sqrt{1 + \tan^2\alpha}} = \frac{2}{\sqrt{5}}.$$

At the point $M_0(1, 1)$ $\partial u/\partial x$ and $\partial u/\partial y$ have the values

$$\left.\frac{\partial u}{\partial x}\right|_{M_0} = \left.\frac{y}{1 + x^2y^2}\right|_{M_0} = \frac{1}{2},$$

$$\left.\frac{\partial u}{\partial y}\right|_{M_0} = \left.\frac{x}{1 + x^2y^2}\right|_{M_0} = \frac{1}{2}.$$

From (24.10) we will then get

$$\left.\frac{\partial u}{\partial x}\right|_{M_0} = \frac{1}{2}\frac{1}{\sqrt{5}} + \frac{1}{2}\frac{2}{\sqrt{5}} = \frac{3}{2\sqrt{5}}. \quad ►$$

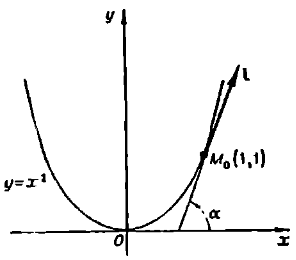


Fig. 24.3

Remark. Formula (24.9) to compute the derivative in the direction of l at a given point M_0 remains valid even when the point M tends to M_0 along the curve for which l is the tangent at M_0.

Example. Find the derivative of the scalar field $u = \ln(xy + yz + zx)$ at the point $M_0(0, 1, 1)$ in the direction of the circle

$$x = \cos t, \quad y = \sin t, \quad z = 1, \quad 0 \leqslant t \leqslant 2\pi.$$

◄ The vector equation of the circle has the form

$$\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + 1\mathbf{k}.$$

We find the vector τ tangent to it at any point M

$$\tau = \frac{d\mathbf{r}}{dt} = -\sin t\,\mathbf{i} + \cos t\,\mathbf{j}.$$

Corresponding to $M_0(0, 1, 1)$ is the value of the parameter $t = \pi/2$. We will thus have

$$\tau|_{M_0} = -\sin\frac{\pi}{2}\mathbf{i} + \cos\frac{\pi}{2}\mathbf{j} = -1\mathbf{i}.$$

It follows that at M_0 $\cos\alpha = -1$, $\cos\beta = 0$, and $\cos\gamma = 0$.

At $M_0(0, 1, 1)$ the partial derivatives of the scalar field are

$$\left.\frac{\partial u}{\partial x}\right|_{M_0} = \left.\frac{y + z}{xy + yz + zx}\right|_{M_0} = 2,$$

$$\left.\frac{\partial u}{\partial y}\right|_{M_0} = \left.\frac{x + z}{xy + yz + zx}\right|_{M_0} = 1,$$

$$\left.\frac{\partial u}{\partial z}\right|_{M_0} = \left.\frac{y + x}{xy + yz + zx}\right|_{M_0} = 1.$$

And so the derivative will be

$$\left.\frac{\partial u}{\partial l}\right|_{M_0} = \left.\frac{\partial u}{\partial \tau}\right|_{M_0} = 2(-1) + 1 \cdot 0 + 1 \cdot 0 = -2. \quad ►$$

24.2. Gradient of a Scalar Field

Consider a scalar field defined by the scalar function

$$u = f(x, y, z),$$

where f is supposed to be differentiable.

Definition. The *gradient* of a scalar field u at a given point M is the vector denoted by grad u and given by

$$\operatorname{grad} u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}. \tag{24.11}$$

It is quite obvious that the vector is dependent both on f and on the position of M.

Let $\mathbf{l}^0$ be a unit vector in the direction of $\mathbf{l}$, i.e.,

$$\mathbf{l}^0 = \frac{\mathbf{l}}{|\mathbf{l}|} = \cos\alpha\,\mathbf{i} + \cos\beta\,\mathbf{j} + \cos\gamma\,\mathbf{k}. \tag{24.12}$$

Equation (24.9) can then be rewritten as

$$\frac{\partial u}{\partial l} = (\operatorname{grad} u, \mathbf{l}^0), \tag{24.13}$$

i.e., the derivative of u in the direction of $\mathbf{l}$ is equal to the scalar product of the gradient of $u(M)$ and $\mathbf{l}^0$.

Basic properties of the gradient are covered by the following theorems:

Theorem 24.1. *The gradient of a scalar field is perpendicular to level surfaces (or level curves if the field is plane).*

◄ Through an arbitrary point M we draw the level surface u = const and choose on this surface a smooth curve L passing through M (Fig. 24.4). Let $\mathbf{l}$ be the vector tangent to L at M.

According to the second remark in Sec. 24.1 we will have

$$\frac{\partial u}{\partial l} = \lim_{\substack{\Delta l \to 0 \\ (M_1 \to M \text{ along } L)}} \frac{u(M_1) - u(M)}{\Delta l} = 0,$$

since on the level surface $u(M) = u(M_1)$ for any point $M_1 \in L$.

On the other hand, $\partial u/\partial l = (\mathbf{grad}\ u, \mathbf{l}^0)$. Therefore, $(\mathbf{grad}\ u, \mathbf{l}^0) = 0$, which suggests that $\mathbf{grad}\, u$ and $\mathbf{l}^0$ are orthogonal vectors, i.e.,

$$\mathbf{grad}\, u \perp \mathbf{l}^0. \tag{24.14}$$

In summary, the vector $\mathbf{grad}\, u$ is orthogonal to any tangent to the level surface at the point M, and hence it is orthogonal to the surface itself at M. ►

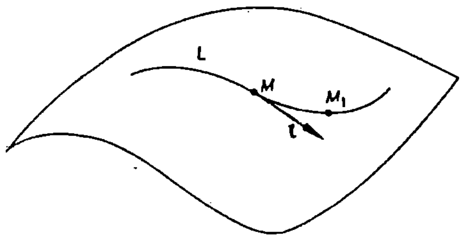


Fig. 24.4

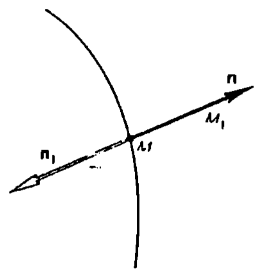


Fig. 24.5

Theorem 24.2 *The gradient points in the direction of increase of the field function.*

◄ From Theorem 24.1 we know already that the gradient points along the normal to the level surface. But the normal to the level surface can be oriented in either of the two directions, i.e., of increase or decrease of the function $u(M)$.

We denote by $\mathbf{n}$ the normal to the level surface oriented in the direction of the growth of $u(M)$ and find the derivative of u along this normal (Fig. 24.5).

$$\frac{\partial u}{\partial n} = \lim_{\substack{\Delta l \to 0 \\ (M_1 \to M \text{ along } \mathbf{n})}} \frac{u(M_1) - u(M)}{\Delta l} = 0.$$

Since, as stated, $u(M_1) > u(M)$, we have $u(M_1) - u(M) > 0$, and therefore

$$\frac{\partial u}{\partial n} = (\mathbf{grad}\, u, \mathbf{n}^0) \geqslant 0,$$

i.e., the projection of $\mathbf{grad}\, u$ on $\mathbf{n}$ is nonnegative.

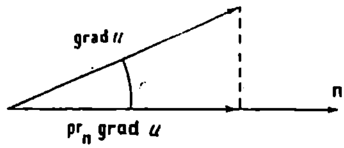


Fig. 24.6

This suggests that **grad** u points in the same direction as the normal **n** we have chosen, i.e., in the direction of growth of $u(M)$ (Fig. 24.6). ►

Theorem 24.3. *The length of a gradient is equal to the largest directional derivative at a given point of the field, i.e.,*

$$\max \frac{\partial u}{\partial l} = |\mathbf{grad}\, u| = \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2}. \quad (24.15)$$

Here $\max(\partial u/\partial l)$ is taken in all the directions possible at a given point M of the field.

◄ We have

$$\frac{\partial u}{\partial l} = (\mathbf{grad}\, u, \mathbf{l}^0) = |\mathbf{grad}\, u| \cdot 1 \cdot \cos\varphi,$$

where φ is the angle between **l** and **grad** u. Since the maximum value of $\cos\varphi$ is 1, then the largest value of $\partial u/\partial x$ is exactly $|\mathbf{grad}\, u|$. ►

Example. At the point $M_0(1, 1, 1)$ it is required to find the direction of the greatest variation of the scalar field

$$u = xy + yz + zx,$$

and also the value of that greatest variation at the point.

◄ The direction of the greatest variation of a scalar field is indicated by **grad** $u(M)$. We thus have

$$\mathbf{grad}\, u(M) = (y + z)\mathbf{i} + (x + z)\mathbf{j} + (y + x)\mathbf{k},$$

so that $\mathbf{grad}\, u(M_0) = 2\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$.

This vector determines the direction of the largest growth of the field at $M_0(1, 1, 1)$. The magnitude of the largest variation of the field at the point is

$$\max \frac{\partial u}{\partial l} = \left|\mathbf{grad}\, u(M_0)\right| = 2\sqrt{3}. \quad ►$$

Invariant definition for the gradient. Quantities that are independent of the choice of coordinate system and are characteristic of some properties of the object under consideration are called *invariants* of the objects. For

example, the length of a curve is an invariant, the angle between the tangent of a curve and the x-axis is not an invariant.

On the basis of the three properties of the gradients proved above we can give the following invariant definition of the gradient of a scalar field.

The *gradient of a scalar field* is a vector pointing along the normal to the level surface in the direction of growth of the field function. The magnitude of the gradient is equal to the largest directional derivative (at the given point).

Consequently, the magnitude and direction of the gradient characterizes the *rate of growth of the field.* The invariant form of the gradient is

$$\mathbf{grad}\, u = |\mathbf{grad}\, u|\, \mathbf{n}^0,$$

where $\mathbf{n}^0$ points in the direction of growth of the field. But $|\mathbf{grad}\, u| = \partial u/\partial n$. Therefore,

$$\mathbf{grad}\, u = \frac{\partial u}{\partial n}\, \mathbf{n}^0. \tag{24.16}$$

Example. Find the gradient of the distance

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}, \tag{24.17}$$

where $M(x, y, z)$ is the field point in question, $M_0(x_0, y_0, z_0)$ is some fixed point of the field.

◄ Clearly, level surfaces of the scalar field (24.17) are spheres with centre at $M_0(x_0, y_0, z_0)$. Let us determine the gradient:

$$\begin{aligned}\mathbf{grad}\, r &= \frac{\partial r}{\partial x}\mathbf{i} + \frac{\partial r}{\partial y}\mathbf{j} + \frac{\partial r}{\partial z}\mathbf{k}\\ &= \frac{(x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}}{\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}} = \mathbf{r}^0\end{aligned}$$

where $\mathbf{r}^0$ is the unit vector of the direction $\overrightarrow{M_0M}$. Thus,

$$\mathbf{grad}\, r = \mathbf{r}^0. \quad ► \tag{24.18}$$

Properties of gradients.

(1) $\mathbf{grad}\, Cu(M) = C\, \mathbf{grad}\, u(M),$ (24.19)

where C is a numerical constant.

(2) $\mathbf{grad}\,(u + v) = \mathbf{grad}\, u + \mathbf{grad}\, v.$ (24.20)

Formulas (24.19) and (24.20) follow directly from the definition of the gradient and the properties of derivatives. They show that the gradient is a vector defined in differentiable scalar fields.

(3) $\mathbf{grad}\,(uv) = v\, \mathbf{grad}\, u + u\, \mathbf{grad}\, v.$ (24.21)

◄ From the rule of differentiation of a product

$$\operatorname{grad}(uv) = \frac{\partial(uv)}{\partial x}\mathbf{i} + \frac{\partial(uv)}{\partial y}\mathbf{j} + \frac{\partial(uv)}{\partial z}\mathbf{k}$$

$$= \left(v\frac{\partial u}{\partial x} + u\frac{\partial v}{\partial x}\right)\mathbf{i} + \left(v\frac{\partial u}{\partial y} + u\frac{\partial v}{\partial y}\right)\mathbf{j} + \left(v\frac{\partial u}{\partial z} + u\frac{\partial v}{\partial z}\right)\mathbf{k}$$

$$= v\left(\frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}\right) + u\left(\frac{\partial v}{\partial x}\mathbf{i} + \frac{\partial v}{\partial y}\mathbf{j} + \frac{\partial v}{\partial z}\mathbf{k}\right)$$

$$= v \operatorname{grad} u + u \operatorname{grad} v. \quad ►$$

(4) $$\operatorname{grad}\left(\frac{u}{v}\right) = \frac{v \operatorname{grad} u - u \operatorname{grad} v}{v^2}, \quad v \neq 0. \tag{24.22}$$

(5) $$\operatorname{grad} F(u) = F'(u) \operatorname{grad} u. \tag{24.23}$$

◄ By the definition of the gradient we have

$$\operatorname{grad} F(u) = \frac{\partial F(u)}{\partial x}\mathbf{i} + \frac{\partial F(u)}{\partial y}\mathbf{j} + \frac{\partial F(u)}{\partial z}\mathbf{k}.$$

To all the terms on the right we can apply the rule of differentiation of a composite function. We obtain

$$\operatorname{grad} F(u) = F'(u)\frac{\partial u}{\partial x}\mathbf{i} + F'(u)\frac{\partial u}{\partial y}\mathbf{j} + F'(u)\frac{\partial u}{\partial z}\mathbf{k}$$

$$= F'(u) \operatorname{grad} u. \quad ►$$

Specifically

$$\operatorname{grad} F(r) = F'(r)\mathbf{r}^0. \tag{24.24}$$

Formula (24.24) follows from (24.23) and the formula $\operatorname{grad} r = \mathbf{r}^0$.

Examples. (1). Find the derivative in the direction of the radius vector $\mathbf{r}$ of the function $u = \sin r$, where $r = |\mathbf{r}|$.

◄ From (24.13) we have

$$\frac{\partial u}{\partial r} = (\operatorname{grad} \sin r, \mathbf{r}^0). \tag{24.25}$$

But by (24.24)

$$\operatorname{grad} \sin r = \mathbf{r}^0 \cos r.$$

Substituting this expression into the right-hand side of (24.25) we will have

$$\frac{\partial u}{\partial r} = (\mathbf{r}^0 \cos r, \mathbf{r}^0) = \cos r. \quad ►$$

(2) Consider the plane scalar field

$$u = r_1 + r_2, \tag{24.26}$$

where r_1 and r_2 are the distances of a point $P(x, y)$ on a plane to two fixed points F_1 and F_2 on the plane. Prove that a beam of light emitted from one focus of the ellipse will get to the other focus.

◄ The level curves of (24.26) are

$$r_1 + r_2 = 2a, \quad a > 0. \tag{24.27}$$

These are the equations of ellipses with foci at points F_1 and F_2.

According to (24.18) we have

$$\operatorname{grad}(r_1 + r_2) = \mathbf{r}_1^0 + \mathbf{r}_2^0,$$

i.e., the gradient is equal to the diagonal $\overrightarrow{PQ}$ of the rhombus constructed on the unit vectors $\mathbf{r}_1^0$ and $\mathbf{r}_2^0$ of radius vectors drawn from F_1 and F_2 to

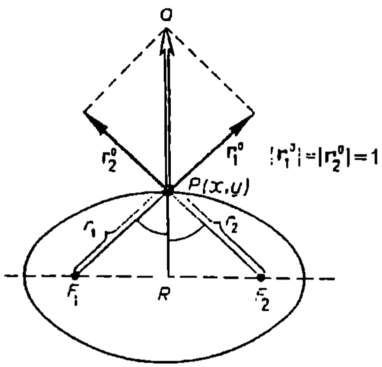


Fig. 24.7

$P(x, y)$ (Fig. 24.7). By Theorem 24.1 the gradient $\overrightarrow{PQ}$ is perpendicular to the ellipse (24.27) at $P(x, y)$. Accordingly, the normal to the ellipse at some point on it halves the angle between the radius vectors drawn to this point. The incidence angle equals the reflection angle, and so the beam of light from one focus of the ellipse comes to the other one. ►

24.3 Vector Field. Vector Lines and Their Differential Equations

Definition. If at each point $M(x, y, z)$ of space or part of space the vector quantity

$$\mathbf{a} = \mathbf{a}(M) = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k} \tag{24.28}$$

is defined, it is said that the *vector field* $\mathbf{a}$ is specifed.

To specify a vector field is equivalent to specifying three functions of three variables

$$P(x, y, z), \quad Q(x, y, z), \quad R(x, y, z).$$

Examples of vector fields are a field of force **F**, a field of velocities **v** in the flow of some liquid, and so on.

Geometrically, a vector field can be represented by *vector lines*. The vector line of a vector field **a** is a curve such that a tangent to it at any point M has the same direction as the vector field **a** at that point (Fig. 24.8).

In a field of force vector lines are called *lines of force*, in a field of velocities of the motion of liquid the lines are called *lines of flow*.

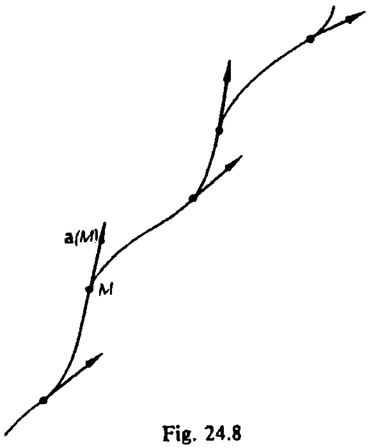


Fig. 24.8

Differential equations of vector lines. Let a vector field be defined by the vector

$$\mathbf{a} = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k},$$

where $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ are continuous functions of the variables x, y, z with bounded partial derivatives of the first order.

It is well known that we may define the vector τ of the tangent to the curve to be the vector $d\mathbf{r}(t)/dt$, where

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}$$

is the radius vector of the running point on the curve, and t is some parameter for the curve under consideration.

It follows from the definition of vector lines that the vectors

$$\mathbf{a} = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k}$$

and

$$\frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k}$$

must be collinear at each point on the vector line. The collinearity condition for vectors is the proportionality of their coordinates, therefore, on the vector line the following relations must be obeyed:

$$\frac{dx}{P(x, y, z)} = \frac{dy}{Q(x, y, z)} = \frac{dz}{R(x, y, z)}. \tag{24.29}$$

Thus, we have obtained a system of differential equations in symmetrical form to determine vector lines.

Suppose that we have found two independent integrals of the system (24.29)

$$\begin{cases} \varphi_1(x, y, z) = C_1, \\ \varphi_2(x, y, z) = C_2. \end{cases} \tag{24.30}$$

The system of two equations (24.30) determines the vector lines as the intersection of two surfaces. By arbitrarily changing the parameters C_1 and C_2, we will obtain a family of vector lines as a family with two parameters (degrees of freedom).

Example. Find the vector lines of the vector field $\mathbf{a} = x\mathbf{i} + y\mathbf{j} + 2z\mathbf{k}$.

◄ We write the differential equations of the vector lines

$$\frac{dx}{x} = \frac{dy}{y} = \frac{dz}{2z},$$

or

$$\frac{dx}{x} = \frac{dy}{y}, \quad \frac{dx}{x} = \frac{dz}{2z}.$$

Integrating this system, we arrive at

$$y = C_1 x,$$
$$z = C_2 x^2,$$

where C_1, C_2 are arbitrary constants.

The intersection of the planes $y = C_1 x$ with the parabolic cylinders $z = C_2 x^2$ gives a two-parameter family of the vector lines of the field (Fig. 24.9). ►

Definition. A vector field is called *plane* if all the vectors $\mathbf{a}$ are parallel to the same plane and if in each plane parallel to the given one the vector field is the same. This can be interpreted as follows:

If we take the given plane (or any one parallel to it) to be the xy-plane,

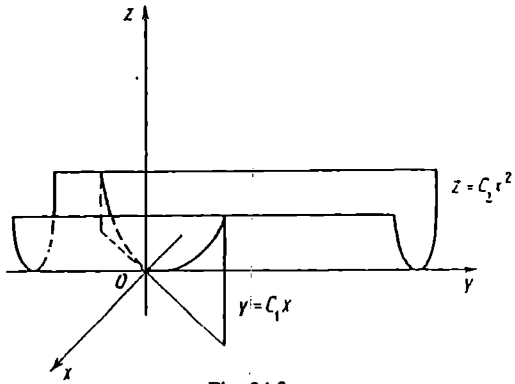


Fig. 24.9

then the vectors of the field will not contain the z-components and the coordinates of the vectors will not be dependent on z

$$\mathbf{a} = P(x, y)\,\mathbf{i} + Q(x, y)\,\mathbf{j}. \tag{24.31}$$

The differential equations of the vector lines of a plane field will have the form

$$\frac{dx}{P(x, y)} = \frac{dy}{Q(x, y)} = \frac{dz}{0},$$

or

$$\begin{cases} \dfrac{dy}{dx} = \dfrac{Q(x, y)}{P(x, y)}, \\ z = \text{const}. \end{cases} \tag{24.32}$$

It follows that the vector lines of the plane field are plane curves that lie in planes parallel to the xy-plane.

Examples. (1) Find the vector lines of the magnetic field of an infinite straight wire.

◄ Suppose that the wire points along the z-axis and carries a current I, i.e., the vector of the current is

$$\mathbf{I} = I\mathbf{k}.$$

The vector of the strength $\mathbf{H}$ of the magnetic field will then be

$$\mathbf{H} = \frac{2}{\varrho^2}[\mathbf{I}, \mathbf{r}], \tag{24.33}$$

where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is the radius vector of the point M under consideration, and ϱ is the distance from the wire axis to M. Expanding the vector

product (24.33) gives

$$\mathbf{H} = -\frac{2Iy}{\varrho^2}\mathbf{i} + \frac{2Ix}{\varrho^2}\mathbf{j}.$$

The differential equations of the vector lines are

$$\frac{dx}{-2Iy/\varrho^2} = \frac{dy}{2Ix/\varrho^2} = \frac{dz}{0}.$$

Hence $z = \text{const}$, $-dx/y = dy/x$, or $x\,dx + y\,dy = 0$. Finally,

$$\begin{cases} x^2 + y^2 = C_1^2, \\ z = C_2, \end{cases}$$

i.e., the vector lines are circles with centres on the z-axis (Fig. 24.10). ►

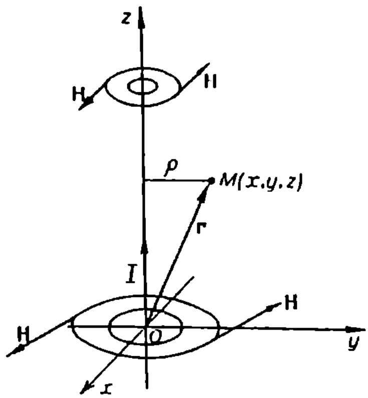


Fig. 24.10

(2) Find the vector lines of the gravitational field formed by an attracting material point of mass m located at the origin of coordinates.

◄ In this case the force $\mathbf{F}$ is determined as

$$\mathbf{F} = -\frac{\gamma m\mathbf{r}}{|\mathbf{r}|^3} = -\frac{\gamma m x\mathbf{i}}{(x^2+y^2+z^2)^{3/2}} - \frac{\gamma m y\mathbf{j}}{(x^2+y^2+z^2)^{3/2}} - \frac{\gamma m z\mathbf{k}}{(x^2+y^2+z^2)^{3/2}}.$$

The vector lines are described by the differential equations

$$\frac{dx}{-\dfrac{\gamma m x}{(x^2+y^2+z^2)^{3/2}}} = \frac{dy}{-\dfrac{\gamma m y}{(x^2+y^2+z^2)^{3/2}}} = \frac{dz}{-\dfrac{\gamma m z}{(x^2+y^2+z^2)^{3/2}}}.$$

Multiplying each fraction here by $-\gamma m/(x^2+y^2+z^2)^{3/2}$ gives

$$\frac{dx}{x}=\frac{dy}{y}=\frac{dz}{z}=\frac{dt}{t}.$$

We have here equated each of the fractions to dt/t.

We can now readily obtain the equations of the vector lines in parametric form

$$x=C_1t,\quad y=C_2t,\quad z=C_3t.$$

These are rays originating at the origin of coordinates (at the origin the vector field is not defined!).

To isolate one line out of a family of all vector lines, we should specify a point $M_0(x_0, y_0, z_0)$ through which this vector line must pass and determine C_1, C_2, C_3 from the coordinates of the point.

Let M_0 have the coordinates $x_0 = 3$, $y_0 = 5$, $z_0 = 7$. The equation of the vector line passing through $M_0(3, 5, 7)$ can be written as

$$x=3t,\quad y=5t,\quad z=7t.$$

The point M_0 is obtained at $t = 1$. ►

24.4 Vector Flux Through a Surface and Its Properties

Consider at first the special case of the field of velocities $\mathbf{v}$ of liquid flow. In the field we isolate a surface Σ. The *flow* of liquid through Σ is the amount of liquid flowing through Σ in a unit time.

We easily determine this flow if the velocity is constant ($\mathbf{v}$ = const) and the surface Σ is a plane. The liquid flow will then be equal to the volume of the cylindrical body with parallel bases and the generatrix of length $|\mathbf{v}|$, since in a unit time each particle shifts by $\mathbf{v}$ (Fig. 24.11).

The flow will then be

$$\Pi = Sh,$$

where S is the area of the base, $h = \text{pr}_{\mathbf{n}}\,\mathbf{v} = (\mathbf{v}, \mathbf{n}^0)$ is the height of the cylinder, $\mathbf{n}$ is the normal to the base. And so at a constant $\mathbf{v}$ the flow of liquid through a plane Σ will be

$$\Pi = (\mathbf{v}, \mathbf{n}^0)\cdot S. \tag{24.34}$$

If $\mathbf{v}$ varies continuously, and the surface Σ is smooth, then we can break Σ into parts Σ_k ($k = 1, 2, \ldots, n$) so small that we may approximately believe that each Σ_k is a plane surface and $\mathbf{v}$ on it is constant.

Since the flow of liquid through Σ is equal to the sum of flows through all Σ_k, we will derive the following approximate formula for the flow

$$\Pi \simeq \sum_{k=1}^{n} (\mathbf{v}, \mathbf{n}^0)_{P_k} \cdot \Delta\sigma_k, \tag{24.35}$$

where n is the number of parts Σ_k into which the surface Σ is divided, P_k is a point in the kth part, $\Delta\sigma_k$ is the area of Σ_k, $(\mathbf{v}, \mathbf{n}^0)_{P_k}$ is the scalar product, where $\mathbf{v}$ and $\mathbf{n}^0$ are taken at $P_k \in \Sigma_k$ (Fig. 24.12).

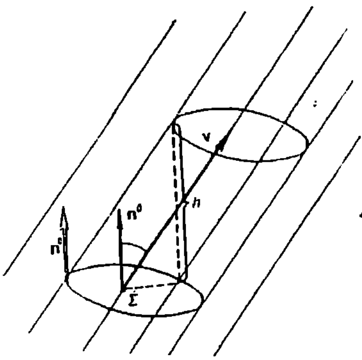


Fig. 24.11

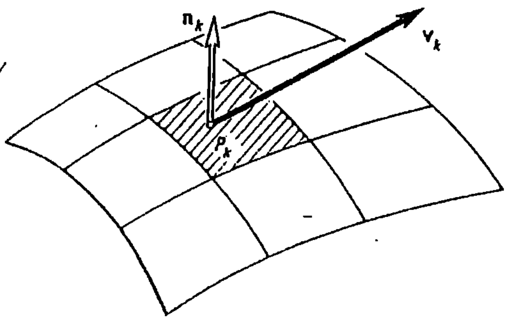


Fig. 24.12

The flow through Σ will then be the limit of (24.35) as the largest of the diameters of Σ_k tends to zero

$$\Pi = \lim_{d \to 0} \sum_{k=1}^{n} (\mathbf{v}, \mathbf{n}^0)_{P_k} \cdot \Delta\sigma_k = \iint_{\Sigma} (\mathbf{v}, \mathbf{n}^0)\, d\sigma, \tag{24.36}$$

where d is the largest of the diameters of Σ_k $(k = 1, 2, \ldots, n)$.

The integral (24.36) defining the flow of liquid is the integral of the scalar function $(\mathbf{v}, \mathbf{n}^0)$ over the surface Σ.

In analogy with the notion of the flow through a surface introduced above we will use the concept of the flux of any vector $\mathbf{a}$ through σ.

Definition. The *flux of vector* $\mathbf{a}$ through a surface Σ is the integral over Σ of the projection of $\mathbf{a}$ on the normal to the surface

$$\Pi = \iint_{\Sigma} (\mathbf{a}, \mathbf{n}^0)\, d\sigma, \tag{24.37}$$

or $\iint_{\Sigma} a_n\, d\sigma$, or $\iint_{\Sigma} (\mathbf{a}, d\boldsymbol{\sigma})$, where $d\boldsymbol{\sigma} = \mathbf{n}^0 d\sigma$. Clearly, integral (24.37)

exists if the vector

$$\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$$

is continuous, i.e., its coordinates $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ are continuous, and the surface Σ is smooth, i.e., it has a continuously varying tangent surface.

Example. Let a field be produced by a point source (electric field) or a point mass (gravitational field) placed at the origin of coordinates. The strength of the field at any point P will be

$$\mathbf{E} = \frac{q}{r^2}\mathbf{r}^0,$$

where q is the charge (mass), $\mathbf{r} = \overrightarrow{OP}$ is the radius vector of P. We desire to find the flux of $\mathbf{E}$ through S_R, a sphere of radius R with centre at the origin.

◄ Since the direction of the normal to the sphere coincides with the direction of the radius vector $\mathbf{r}$, we have $\mathbf{n}^0 = \mathbf{r}^0$, and so

$$(\mathbf{E}, \mathbf{n}^0) = (\mathbf{E}, \mathbf{r}^0) = \frac{q}{r^2}.$$

On the sphere S_R we have $r = R$, so that $(\mathbf{E}, \mathbf{n}^0) = q/R^2 = \text{const}$. Therefore, the flux of the vector through S_R is

$$\Pi = \iint\limits_{S_R} (\mathbf{E}, \mathbf{n}^0)\, d\sigma = \frac{q}{R^2} \iint\limits_{S_R} d\sigma = \frac{q}{R^2} 4\pi R^2 = 4\pi q. \quad ►$$

Properties of the flux of a vector across a surface.

(1) *Linearity*

$$\iint\limits_{\Sigma} (\lambda\mathbf{a} + \mu\mathbf{b}, \mathbf{n}^0)\, d\sigma = \lambda \iint\limits_{\Sigma} (\mathbf{a}, \mathbf{n}^0)\, d\sigma + \mu \iint\limits_{\Sigma} (\mathbf{b}, \mathbf{n}^0)\, d\sigma, \tag{24.38}$$

where λ and μ are constants.

(2) *Additivity.* If a surface Σ is divided into Σ_1 and Σ_2 by a piecewise smooth curve, then the flux through Σ will be equal to the sum of the fluxes through Σ_1 and Σ_2

$$\iint\limits_{\Sigma} (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iint\limits_{\Sigma_1} (\mathbf{a}, \mathbf{n}^0)\, d\sigma + \iint\limits_{\Sigma_2} (\mathbf{a}, \mathbf{n}^0)\, d\sigma. \tag{24.39}$$

This property enables us to generalize the concept of the flux to piecewise smooth surfaces.

(3) *Dependence of the flux on surface orientation* (i.e., on the orientation of the normal to a surface). So far we have not discussed the choice of the normal **n** to the surface Σ. Let us now look at this question.

Take, say, a cylindrical surface. If at a point M on the surface we choose a normal **n** and then move continuously over the surface along any path without venturing beyond the boundary of the surface, then we will always return to M with the previous orientation of **n**.

To obtain the opposite of **n** we will have to go over onto the other side of the surface. Surfaces of this kind are known as *two-sided* or *orientable*. Among them are the plane, the sphere, the ellipsoid, the surface of a cube, ans so forth.

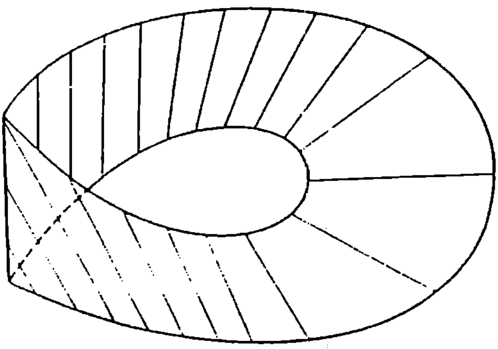


Fig. 24.13

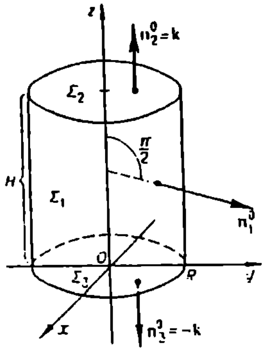


Fig. 24.14

There also exist *one-sided*, or *nonorientable*, surfaces, on which it is impossible to select a certain orientation of the normal. An example of such surfaces is the *Möbius strip* (Fig. 24.13). If at some point in the Möbius strip we take a normal **n**, there exists a *path* (e.g. the central line of the strip) such that if we move a point continuously along it we will arrive at the same point on the surface but with the opposite orientation of the normal.

We introduce the concept of the flux for two-sided surfaces only. We will assume that if at one point of such a surface we have already chosen the direction of the normal, then at another point we take the normal that results if we continuously move the initial point to that point (without crossing the boundary of the surface).

In particular, on a closed surface we take at all points either the external or the internal normal (an internal normal points inside the body bounded by the closed surface).

We denote by Σ^+ the side of Σ on which we take $\mathbf{n}$, and by Σ^- the side of Σ on which we take $-\mathbf{n}$. We then obtain ($\mathbf{n}^0_- = -\mathbf{n}^0_+$)

$$\iint\limits_{\Sigma^-}(\mathbf{a}, \mathbf{n}^0_-)\,d\sigma = -\iint\limits_{\Sigma^+}(\mathbf{a}, \mathbf{n}^0_+)\,d\sigma. \tag{24.40}$$

Consequently, when the orientation of the surface is changed (i.e., the direction of $\mathbf{n}^0$ to Σ is changed) the flux of the vector is reversed.

Example. Calculate the flux of the radius vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ through the surface of a right circular cylinder of height H, base radius R and the z-axis as its axis.

◄ The surface Σ consists of three parts: side surface Σ_1, upper base Σ_2 and lower base Σ_3 of the cylinder.

According to the additivity property

$$\Pi = \Pi_1 + \Pi_2 + \Pi_3,$$

where Π_1, Π_2, Π_3 are the fluxes of the field through Σ_1, Σ_2, and Σ_3, respectively.

On the side surface of the cylinder the external normal $\mathbf{n}^0_1$ is parallel to the xy-plane, and therefore (see Fig. 24.14)

$$(\mathbf{r}, \mathbf{n}^0_1) = \mathrm{pr}_{\mathbf{n}^0_1}\mathbf{r} = R.$$

Hence,

$$\Pi_1 = \iint\limits_{\Sigma_1}(\mathbf{r}, \mathbf{n}^0)\,d\sigma = R\iint\limits_{\Sigma_1}d\sigma = R\,2\pi RH = 2\pi R^2H.$$

On the upper base Σ_2 the normal $\mathbf{n}^0_2$ is parallel to the z-axis, and therefore we can put $\mathbf{n}^0_2 = \mathbf{k}$. Then

$$(\mathbf{r}, \mathbf{n}^0_2) = (\mathbf{r}, \mathbf{k}) = \mathrm{pr}_{\mathbf{k}}\mathbf{r} = H,$$

so that

$$\Pi_2 = \iint\limits_{\Sigma_2}(\mathbf{r}, \mathbf{n}^0_2)d\sigma = H\iint\limits_{\Sigma_2}d\sigma = \pi R^2H.$$

On the lower base Σ_3 the vector $\mathbf{r}$ is perpendicular to the normal $\mathbf{n}^0_3 = -\mathbf{k}$. Therefore,

$$(\mathbf{r}, \mathbf{n}^0_3) = (\mathbf{r}, -\mathbf{k}) = 0$$

and

$$\Pi_3 = \iint\limits_{\Sigma_3}(\mathbf{r}, \mathbf{n}^0_3)\,d\sigma = 0.$$

The flux will thus be

$$\Pi = \oint\!\!\!\oint_{\Sigma} (\mathbf{r}, \mathbf{n}^0)\, d\sigma = 3\pi R^2 H. \quad \blacktriangleright$$

Here $\oint\!\!\!\oint_{\Sigma}$ stands for a double integral along a closed surface.

24.5 Flux of a Vector Through an Open Surface

We will discuss several methods of calculating the flux of a vector through open surfaces:

Projection on a coordinate plane. A surface S is uniquely projectable on a domain D_{xy} in the xy-plane. We can then specify S by the equation $z = f(x, y)$, and since the surface element $d\sigma$ of this surface is

$$d\sigma = \frac{dx\,dy}{|\cos\gamma|} = \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dx\,dy, \tag{24.41}$$

to determine the flux Π through a side of S means to take a double integral by

$$\Pi = \iint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iint_{D_{xy}} \left.\frac{(\mathbf{a}, \mathbf{n}^0)}{|\cos\gamma|}\right|_{z = f(x, y)} dx\,dy. \tag{24.42}$$

Here the unit vector $\mathbf{n}^0$ of the normal to the chosen side of the surface is found by

$$\mathbf{n}^0 = \pm \frac{\operatorname{grad}[z - f(x, y)]}{|\operatorname{grad}[z - f(x, y)]|} = \pm \frac{-\dfrac{\partial f}{\partial x}\mathbf{i} - \dfrac{\partial f}{\partial y}\mathbf{j} + \mathbf{k}}{\sqrt{1 + \left(\dfrac{\partial f}{\partial x}\right)^2 + \left(\dfrac{\partial f}{\partial y}\right)^2}}. \tag{24.43}$$

We here take the plus sign, if the angle γ between the z-axis and $\mathbf{n}^0$ is acute; if the angle γ is obtuse, then we take the minus sign.

The symbol

$$\left.\frac{(\mathbf{a}, \mathbf{n}^0)}{|\cos\gamma|}\right|_{z = f(x, y)}$$

means that we should substitute $f(x, y)$ for z in the integrand.

Example. Find the flux of the vector $\mathbf{a} = y^2\mathbf{j} + z\mathbf{k}$ through the part of the paraboloid surface $z = x^2 + y^2$ cut by the plane $z = 2$. The normal

is the external one in relation to the region bounded by the paraboloid (Fig. 24.15).

◄ This surface (paraboloid of revolution) is uniquely projectable on the xy-plane to yield the circle D_{xy} with centre at the origin of coordinates and radius $R = \sqrt{2}$. We find the unit vector $\mathbf{n}^0$ of the normal to S:

$$\mathbf{n}^0 = \pm\frac{\operatorname{grad}(z - x^2 - y^2)}{|\operatorname{grad}(z - x^2 - y^2)|} = \pm\frac{-2x\mathbf{i} - 2y\mathbf{j} + \mathbf{k}}{\sqrt{4x^2 + 4y^2 + 1}}.$$

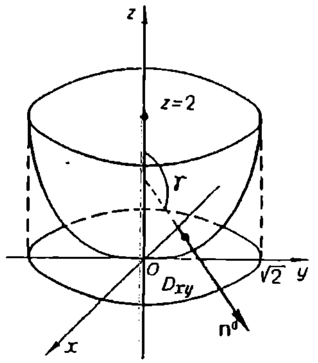


Fig. 24.15

As assumed, $\mathbf{n}^0$ forms an obtuse angle γ with the z-axis, therefore we should choose the minus sign. Thus,

$$\mathbf{n}^0 = \frac{2x\mathbf{i} + 2y\mathbf{i} - \mathbf{k}}{\sqrt{4x^2 + 4y^2 + 1}}. \tag{24.44}$$

Further,

$$(\mathbf{a}, \mathbf{n}^0) = \frac{2y^3 - z}{\sqrt{4x^2 + 4y^2 + 1}}.$$

Since $\cos\gamma$ is the coefficient at the unit vector $\mathbf{k}$ in (24.44),

$$\cos\gamma = -\frac{1}{\sqrt{4x^2 + 4y^2 + 1}},$$

hence

$$\frac{1}{|\cos\gamma|} = \sqrt{4x^2 + 4y^2 + 1}.$$

Therefore,

$$\left.\frac{(\mathbf{a}, \mathbf{n}^0)}{|\cos\gamma|}\right|_{z=f(x,y)} = \left.\left(\frac{2y^3 - z}{\sqrt{4x^2+4y^2+1}}\sqrt{4x^2+4y^2+1}\right)\right|_{z=x^2+y^2}$$
$$= 2y^3 - x^2 - y^2.$$

From (24.42) the flux will be

$$\Pi = \iint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iint_{D_{xy}} (2y^3 - x^2 - y^2)\, dx\, dy.$$

We now introduce the polar coordinates $x = \varrho\cos\varphi$, $y = \varrho\sin\varphi$, where $0 \leqslant \varrho \leqslant \sqrt{2}$, $0 \leqslant \varphi \leqslant 2\pi$; we will have

$$\Pi = \int_0^{2\pi} d\varphi \int_0^{\sqrt{2}} (2\varrho^4 \sin^3\varphi - \varrho^3)\, d\varrho = -2\pi \left.\frac{\varrho^4}{4}\right|_0^{\sqrt{2}} = -2\pi. \quad ►$$

If a surface S is uniquely projectable on a region D_{yz} of the yz-plane, then it can be given by the equation $x = \varphi(y, z)$. In that case, we have

$$d\sigma = \frac{dy\,dz}{|\cos\alpha|} = \sqrt{1 + \left(\frac{\partial\varphi}{\partial y}\right)^2 + \left(\frac{\partial\varphi}{\partial z}\right)^2}\, dy\, dz, \tag{24.45}$$

$$\Pi = \iint_{D_{yz}} \left.\frac{(\mathbf{a}, \mathbf{n}^0)}{|\cos\alpha|}\right|_{x=\varphi(y,z)} dy\, dz, \tag{24.46}$$

where

$$\mathbf{n}^0 = \pm\frac{\operatorname{grad}[x - \varphi(y, z)]}{|\operatorname{grad}[x - \varphi(x, y)]|} = \pm\frac{\mathbf{i} - \dfrac{\partial\varphi}{\partial y}\mathbf{j} - \dfrac{\partial\varphi}{\partial z}\mathbf{k}}{\sqrt{1 + \left(\dfrac{\partial\varphi}{\partial y}\right)^2 + \left(\dfrac{\partial\varphi}{\partial z}\right)^2}}. \tag{24.47}$$

In the last formula we take the plus sign if the angle α between the x-axis and $\mathbf{n}^0$ is acute; and the minus sign if the angle is obtuse.

Lastly, if S is uniquely projectable on the domain D_{xz} in the xz-plane, then it can be defined by the equation $y = \psi(x, z)$, and then

$$d\sigma = \frac{dx\,dz}{|\cos\beta|} = \sqrt{1 + \left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial z}\right)^2}\, dx\, dz, \tag{24.48}$$

$$\Pi = \iint_{D_{xz}} \left.\frac{(\mathbf{a}, \mathbf{n}^0)}{|\cos\beta|}\right|_{y=\psi(x,z)} dx\, dz, \tag{24.49}$$

where

$$\mathbf{n}^0 = \pm\frac{\operatorname{grad}[y - \psi(x, y)]}{|\operatorname{grad}[y - \psi(x, y)]|} = \pm\frac{-\dfrac{\partial\psi}{\partial x} + \mathbf{j} - \dfrac{\partial\psi}{\partial z}\mathbf{k}}{\sqrt{1 + \left(\dfrac{\partial\psi}{\partial x}\right)^2 + \left(\dfrac{\partial\psi}{\partial z}\right)^2}} \tag{24.50}$$

We here use the plus sign if the angle β between the y-axis and $\mathbf{n}^0$ is acute, and the minus sign if the angle β is obtuse.

Remark. If we want to find the flux of the vector

$$\mathbf{a} = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k}$$

through a surface S defined, say, by the equation

$$z = f(x, y)$$

by projecting on the xy-plane, it is not necessary to determine the unit vector $\mathbf{n}^0$ of the normal. Instead, we can take the vector

$$\mathbf{n} = \pm\operatorname{grad}[z - f(x, y)] = \pm\left(-\frac{\partial f}{\partial x}\mathbf{i} - \frac{\partial f}{\partial y}\mathbf{j} + \mathbf{k}\right).$$

Formula (24.42) will then become

$$\begin{aligned}\Pi &= \iint_S (\mathbf{a}, \mathbf{n}^0)\,d\sigma = \iint_{D_{xy}} (\mathbf{a}, \mathbf{n})\Big|_{z = f(x, y)} dx\,dy \\ &= \pm\iint_{D_{xy}} \Big\{-P[x, y, f(x, y)]\frac{\partial f}{\partial x} - Q[x, y, f(x, y)]\frac{\partial f}{\partial y} \\ &\quad + R[x, y, f(x, y)]\Big\}\,dx\,dy. \end{aligned} \tag{24.51}$$

Similar formulas are derived for fluxes through surfaces given by the equation $x = \varphi(y, z)$ or $y = \psi(x, z)$.

Example. Find the flux of the vector $\mathbf{a} = xz\mathbf{i}$ through the external side of the paraboloid $z = 1 - x^2 - y^2$ bounded by the plane $z = 0$ ($z \geqslant 0$) (Fig. 24.16).

◄ We have

$$\mathbf{n} = \pm\operatorname{grad}[z - 1 + x^2 + y^2] = \pm(2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k}).$$

We take the plus sign, since the angle γ is acute. Then

$$(\mathbf{a}, \mathbf{n})\Big|_{z = f(x, y)} = 2x^2(1 - x^2 - y^2).$$

The flux will be

$$\Pi = \iint\limits_{D_{xy}} 2x^2(1 - x^2 - y^2)\,dx\,dy.$$

Passing to the polar coordinates $x = \varrho\cos\varphi$, $y = \varrho\sin\varphi$, $0 \leqslant \varrho \leqslant 1$, $0 \leqslant \varphi \leqslant 2\pi$, we get

$$\Pi = \int\limits_0^{2\pi} 2\cos^2\varphi\,d\varphi \int\limits_0^1 \varrho^3(1 - \varrho^2)\,d\varrho$$

$$= \int\limits_0^{2\pi} \left(\frac{\varrho^4}{4} - \frac{\varrho^6}{6}\right)\Bigg|_{\varrho=0}^{1} (1 + \cos 2\varphi)\,d\varphi = \frac{\pi}{6}. \blacktriangleright$$

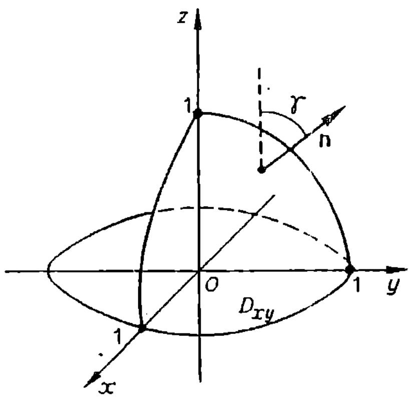


Fig. 24.16

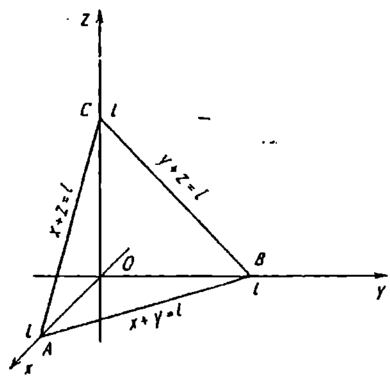


Fig. 24.17

Projecting on three coordinate planes. Let a surface S be uniquely projectable on all the three coordinate planes. We denote by D_{xy}, D_{xz}, and D_{yz} the projections of S of the xy-, xz-, and yz-planes, respectively.

The equation $F(x, y, z) = 0$ of S is then uniquely solvable for each argument, i.e.,

$$x = x(y, z), \quad y = y(x, z), \quad z = z(x, y). \tag{24.52}$$

The flux of the vector

$$\mathbf{a} = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k}$$

through the surface S whose normal unit vector is

$$\mathbf{n}^0 = \mathbf{i}\cos\alpha + \mathbf{j}\cos\beta + \mathbf{k}\cos\gamma$$

can be written as

$$\Pi = \iint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iint_S [P(x, y, z)\cos\alpha$$

$$+ Q(x, y, z)\cos\beta + R(x, y, z)]\, d\sigma. \tag{24.53}$$

It is well known that

$$\begin{cases} d\sigma\cos\alpha = \pm dy\, dz, \\ d\sigma\cos\beta = \pm dx\, dz, \\ d\sigma\cos\gamma = \pm dx\, dy. \end{cases} \tag{24.54}$$

The sign in each of formulas (24.54) is taken to correspond to the sign of $\cos\alpha$, $\cos\beta$, $\cos\gamma$ on S. Substituting (24.52) and (24.54) into (24.53) yields

$$\Pi = \pm \iint_{D_{yz}} P[x(y, z), y, z]\, dy\, dz \pm \iint_{D_{xz}} Q[x, y(x, z), z]\, dx\, dz$$

$$\pm \iint_{D_{xy}} R[x, y, z(x, y)]\, dx\, dy. \tag{24.55}$$

Example. Calculate the flux of the vector field $\mathbf{a} = y\mathbf{i} + z\mathbf{j} + x\mathbf{k}$ through the triangle bounded by the planes $x + y + z = l\ (l > 0)$, $x = 0$, $y = 0$, $z = 0$ (the angle γ is acute) (Fig. 24.17).

◄ We have

$$\mathbf{n}^0 = \frac{\mathbf{grad}\,(x + y + z - l)}{|\mathbf{grad}\,(x + y + z - l)|} = \frac{\mathbf{i} + \mathbf{j} + \mathbf{k}}{\sqrt{3}},$$

so that $\cos\alpha = \cos\beta = \cos\gamma = 1/\sqrt{3} > 0$.

This implies that we take all the integrals in (24.55) with the plus sign. Assuming in (24.55) $P = y$, $Q = z$ and $R = x$, we obtain

$$\Pi = \iint_{D_{yz}} y\, dy\, dz + \iint_{D_{xz}} z\, dx\, dz + \iint_{D_{xy}} x\, dx\, dy. \tag{24.56}$$

The domain D_{yz} is the triangle BOC in the yz-plane, the side BC being given by $y + z = l$, $0 \leqslant y \leqslant l$.

We take the first integral on the right-hand side of (24.56)

$$I_1 = \iint_{D_{yz}} y\,dy\,dz = \int_0^l y\,dy \int_0^{l-y} dz = \int_0^l (l-y)\,y\,dy$$

$$= \left(\frac{ly^2}{2} - \frac{y^3}{3}\right)\Bigg|_0^l = \frac{l^3}{6}.$$

Likewise,

$$I_2 = \iint_{D_{xz}} z\,dx\,dz = \frac{l^3}{6}, \quad I_3 = \iint_{D_{xy}} x\,dx\,dy = \frac{l^3}{6}.$$

And so the flux will be

$$\Pi = I_1 + I_2 + I_3 = \frac{l^3}{2}. \blacktriangleright$$

Using curvilinear coordinates on a surface. If a surface S is part of a circular cylinder or sphere, it is convenient to seek the flux by introducing curvilinear coordinates of the surface without projecting on the coordinate planes.

(a) Let S be part of a circular cylinder

$$x^2 + y^2 = R^2 \tag{24.57}$$

bounded by the surfaces $z = f_1(x, y)$ and $z = f_2(x, y)$, where $f_1(x, y) \leqslant f_2(x, y)$. Putting $x = R\cos\varphi$, $y = R\sin\varphi$, $z = z$, we will have

$$f_1(R\cos\varphi, R\sin\varphi) \leqslant z \leqslant f_2(R\cos\varphi, R\sin\varphi), \ 0 \leqslant \varphi \leqslant 2\pi.$$

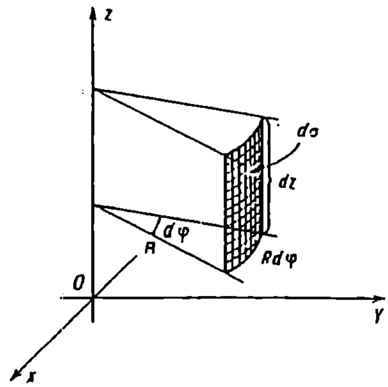


Fig. 24.18

The surface element will be (Fig. 24.18)

$$d\sigma = R\,d\varphi\,dz.$$

The flux of vector a through the external side of S is given by

$$\Pi = R\int_0^{2\pi} d\varphi \int_{f_1(R\cos\varphi,\ R\sin\varphi)}^{f_2(R\cos\varphi,\ R\sin\varphi)} (\mathbf{a}, \mathbf{n}^0)\,dz, \tag{24.58}$$

where

$$\mathbf{n}^0 = \frac{\operatorname{grad}(x^2+y^2-R^2)}{|\operatorname{grad}(x^2+y^2-R^2)|} = \frac{2x\mathbf{i}+2y\mathbf{j}}{\sqrt{4x^2+4y^2}} = \frac{x\mathbf{i}+y\mathbf{j}}{R}. \tag{24.59}$$

Example. Find the flux of the vector $\mathbf{a} = y\mathbf{i} + x\mathbf{j} - e^{yz}\,\mathbf{k}$ through the external side of the cylindrical surface $x^2 + y^2 = 4$ bounded by the planes $z = 0$ and $x + y + z = 4$.

◄ We have $R = 2$, $f_1(x, y) = 0$, $f_2(x, y) = 4 - x - y$. Since

$$\mathbf{n}^0 = \frac{x\mathbf{i}+y\mathbf{j}}{R} = \frac{x}{2}\mathbf{i} + \frac{y}{2}\mathbf{j},$$

the scalar product $(\mathbf{a}, \mathbf{n}^0)$ on the cylinder will be (assuming $x = 2\cos\varphi$, $y = 2\sin\varphi$, $z = z$)

$$(\mathbf{a}, \mathbf{n}^0) = xy = 4\cos\varphi\,\sin\varphi.$$

From (24.58) we have

$$\Pi = 2\int_0^{2\pi} d\varphi \int_0^{4-2\cos\varphi-2\sin\varphi} (\mathbf{a}, \mathbf{n}^0)\,dz$$

$$= 8\int_0^{2\pi} \cos\varphi\,\sin\varphi\,d\varphi \int_0^{4-2\cos\varphi-2\sin\varphi} dz$$

$$= 8\int_0^{2\pi} (4\cos\varphi\sin\varphi - 2\cos^2\varphi\sin\varphi - 2\sin^2\varphi\cos\varphi)\,d\varphi = 0. \ ►$$

(b) Let S be part of the sphere

$$x^2 + y^2 + z^2 = R^2 \tag{24.60}$$

bounded by conic surfaces whose equations in spherical coordinates have the form

$$\theta = \theta_1(\varphi) \quad \text{and} \quad \theta = \theta_2(\varphi)$$

and half-planes $\varphi = \varphi_1$ and $\varphi = \varphi_2$.

For points of the sphere we have

$$x = R\cos\varphi\sin\theta, \quad y = R\sin\varphi\sin\theta, \quad z = R\cos\theta,$$

where $\varphi_1 \leqslant \varphi \leqslant \varphi_2$, $\theta_1 \leqslant \theta \leqslant \theta_2$. The surface element (Fig. 24.19) will then be

$$d\sigma = R^2 \sin\theta \, d\theta \, d\varphi.$$

The flux of the vector field a through the external part S of the sphere is

$$\Pi = R^2 \int_{\varphi_1}^{\varphi_2} d\varphi \int_{\theta_1(\varphi)}^{\theta_2(\varphi)} \sin\theta \, (\mathbf{a}, \mathbf{n}^0) \, d\theta, \tag{24.61}$$

where

$$\mathbf{n}^0 = \frac{\operatorname{grad}(x^2 + y^2 + z^2 - R^2)}{|\operatorname{grad}(x^2 + y^2 + z^2 - R^2)|} = \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{R}. \tag{24.62}$$

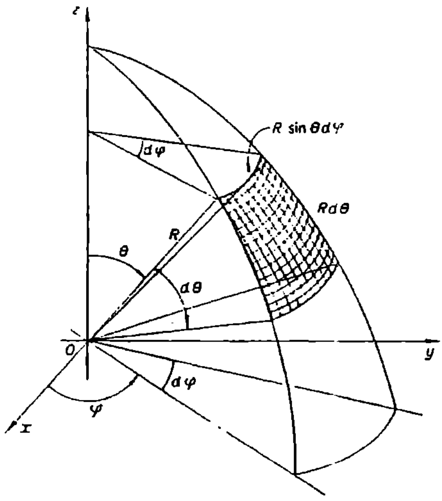


Fig. 24.19

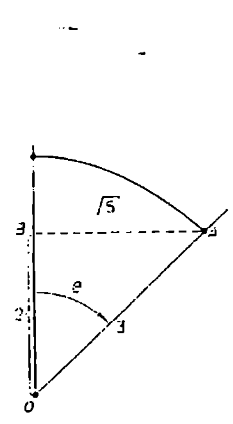


Fig. 24.20

Example. Find the flux of the vector $\mathbf{a} = xz\mathbf{i} + yz\mathbf{j} + z^2\mathbf{k}$ through the external part of the sphere $x^2 + y^2 + z^2 = 9$ cut off by the plane $z = 2$ ($z \geqslant 2$) (Fig. 24.20).

◄ We have $R = 3$, $0 \leqslant \varphi \leqslant 2\pi$; $0 \leqslant \theta \leqslant \cos^{-1}(2/3)$, $\mathbf{n}^0 = (x\mathbf{i} + y\mathbf{j} + z\mathbf{k})/3$, $d\sigma = 9\sin\theta \, d\theta \, d\varphi$. We put $x = 3\sin\theta\cos\varphi$, $y = 3\sin\theta\sin\varphi$, $z = 3\cos\theta$.

The scalar product $(\mathbf{a}, \mathbf{n}^0)$ will then be expressed as follows:

$$(\mathbf{a}, \mathbf{n}^0) = \frac{x^2 z + y^2 z + z^3}{3} = \frac{z(x^2 + y^2 + z^2)}{3} = 9\cos\theta.$$

By (24.61),

$$\Pi = 9\int_0^{2\pi} d\varphi \int_0^{\cos^{-1}\frac{2}{3}} 9\cos\theta \sin\theta\, d\theta = 18\cdot\pi\cdot 9\,\frac{\sin^2\theta}{2}\Bigg|_{\theta=0}^{\cos^{-1}\frac{2}{3}} = 45\pi.$$

Here we have used the formula

$$\sin^2(\cos^{-1}\alpha) = 1 - \cos^2(\cos^{-1}\alpha) = 1 - \alpha^2.$$ ►

24.6 Flux of a Vector Through a Closed Surface. Ostrogradsky-Gauss Formula

Theorem 24.4. *If in a certain space domain $G \subset R^3$ the coordinates of the vector*

$$\mathbf{a} = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k}$$

are continuous and have continuous partial derivatives $\frac{\partial P}{\partial x}$, $\frac{\partial Q}{\partial y}$, $\frac{\partial R}{\partial z}$, *then the flux of* $\mathbf{a}$ *through any closed piecewise smooth surface S lying in a domain G is equal to the triple integral of* $\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$ *over the domain V bounded by the surface S*

$$\Pi = \oiint_S (\mathbf{a}, \mathbf{n}^0)\,d\sigma = \iiint_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dv. \quad (24.63)$$

This is the *Ostrogradsky* (or *Gauss*) *formula*. Here $\mathbf{n}^0$ is the unit vector of the external normal to the surface, and $\oiint_S$ denotes the flux through the closed surface S.

◄ (1) Consider at first the case of $\mathbf{a}$ having only one component

$$\mathbf{a} = R(x, y, z)\,\mathbf{k}$$

and assume that the smooth surface S is intersected at no more than two points by each straight line parallel to the z-axis. This divides S into two parts S_1 and S_2 that are uniquely projectable onto some domain D of the xy-plane (Fig. 24.21).

The external normal to S_2 forms an acute angle γ with the z-axis, and the external normal to S_1 forms an obtuse angle with the z-axis. Therefore, $\cos\gamma = (\mathbf{n}^0, \mathbf{k}) > 0$ on S_2 and $\cos\gamma < 0$ on S_1, so that on S_2 we have $\cos\gamma = |\cos\gamma|$, and on S_1 $\cos\gamma = -|\cos\gamma|$.

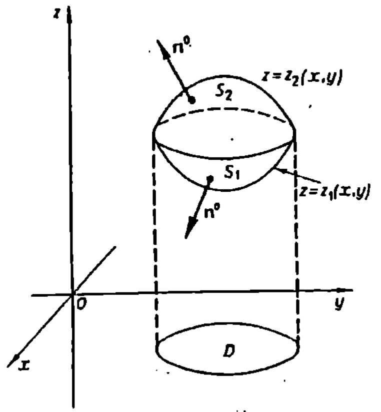


Fig. 24.21

The flux being additive, we have

$$\Pi = \oint\!\!\oint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \oint\!\!\oint_S (R\mathbf{k}, \mathbf{n}^0)\, d\sigma = \oint\!\!\oint_S R \cos\gamma\, d\sigma$$

$$= \iint_{S_2} R \cos\gamma\, d\sigma + \iint_{S_1} R \cos\gamma\, d\sigma$$

$$= \iint_{S_2} R(x, y, z)\, |\cos\gamma|\, d\sigma - \iint_{S_1} R(x, y, z)\, |\cos\gamma|\, d\sigma. \quad (24.64)$$

Let $d\sigma$ be a surface element of S. Then

$$|\cos\gamma|\, d\sigma = ds,$$

where ds is a surface element of the domain D.

We reduce integrals over the surface to double integrals over the domain D (in the xy-plane) onto which the surfaces S_2 and S_1 are projected. Suppose that S_2 is described by the equation $z = z_2(x, y)$, and S_1 by the equation $z = z_1(x, y)$. Then

$$\oint\!\!\oint_S (R\mathbf{k}, \mathbf{n}^0)\, d\sigma = \iint_D R(x, y, z_2(x, y))\, ds$$

$$- \iint_D R(x, y, z_1(x, y))\, ds = \iint_D \big[R(x, y, z_2(x, y))$$

$$- R(x, y, z_1(x, y))\big]\, ds. \quad (24.65)$$

Since the increment of a continuously differentialbe function can be represented as an integral of its derivative

$$F(z_2) - F(z_1) = \int_{z_1}^{z_2} F'(z)\,dz,$$

we will have

$$R(x, y, z_2(x, y)) - R(x, y, z_1(x, y)) = \int_{z_1(x, y)}^{z_2(x, y)} \frac{\partial R(x, y, z)}{\partial z}\,dz.$$

From this and (24.65), we obtain

$$\oiint_S (R\mathbf{k}, \mathbf{n}^0)\,d\sigma = \iint_D ds \int_{z_1(x, y)}^{z_2(x, y)} \frac{\partial R(x, y, z)}{\partial z}\,dz.$$

i. e.,

$$\oiint_S (R\mathbf{k}, \mathbf{n}^0)\,d\sigma = \iiint_V \frac{\partial R}{\partial z}\,dv. \qquad (24.66)$$

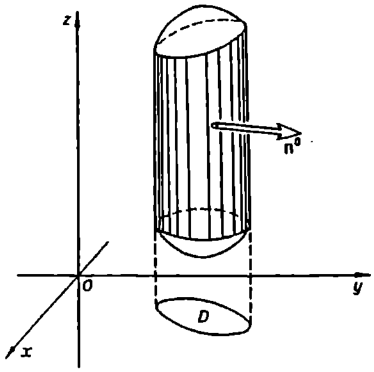


Fig. 24.22

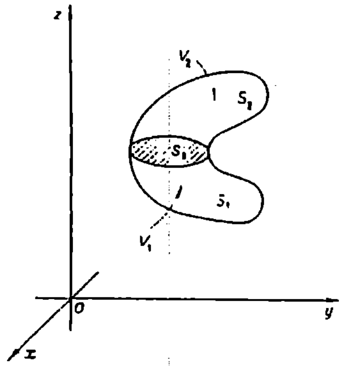


Fig. 24.23

If the surface S contains some part of a cylindrical surface with a generatrix parallel to the z-axis (Fig. 24.22), then on that part of the surface $(R\mathbf{k}, \mathbf{n}^0) = 0$ and integral $\iint (R\mathbf{k}, \mathbf{n}^0)\,d\sigma$ over it is zero. Therefore, formula (24.66) remains valid for surfaces that include such cylindrical parts.

(2) Formula (24.66) can be generalized to such S that are intersected by a vertical straight line at more than two points (Fig. 24.23).

We divide the domain V into parts whose surfaces are intersected by vertical straight lines at more than two points.

Let S_s be the section surface, S_1 and S_2 be those parts of S into which the surface is divided by the section S_s, and let V_1 and V_2 be the corresponding parts of V bounded by $S_1 \cup S_s^+$ and $S_2 \cup S_s^-$.

Then for V_1 we will have

$$\oiint\limits_{S_1 \cup S_s^+} (R\mathbf{k}, \mathbf{n}^0)\, d\sigma = \iiint\limits_{V_1} \frac{\partial R}{\partial z}\, dv,$$

and for V_2

$$\oiint\limits_{S_1 \cup S_s^-} (R\mathbf{k}, \mathbf{n}^0)\, d\sigma = \iiint\limits_{V_2} \frac{\partial R}{\partial z}\, dv.$$

Here the symbol S_s^+means that the normal to S_s is directed upwards (forms an acute angle with the z-axis), and S_s^-means that it is directed downwards (forms an obtuse angle with the z-axis).

Combining these relations and making use of the additivity property of the flux and triple integrals, we obtain

$$\oiint\limits_{S} (R\mathbf{k}, \mathbf{n}^0)\, d\sigma = \iiint\limits_{V} \frac{\partial R}{\partial z}\, dv,$$

since the integrals over S_s cancel out.

(3) Lastly, consider the vector

$$\mathbf{a} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}.$$

For each component $P\mathbf{i}$, $Q\mathbf{j}$ and $R\mathbf{k}$ we can write a formula similar to (24.66) (because all the components are equivalent):

$$\oiint\limits_{S} (P\mathbf{i}, \mathbf{n}^0)\, d\sigma = \iiint\limits_{V} \frac{\partial P}{\partial x}\, dv,$$

$$\oiint\limits_{S} (Q\mathbf{j}, \mathbf{n}^0)\, d\sigma = \iiint\limits_{V} \frac{\partial Q}{\partial y}\, dv,$$

$$\oiint\limits_{S} (R\mathbf{k}, \mathbf{n}^0)\, d\sigma = \iiint\limits_{V} \frac{\partial R}{\partial z}\, dv.$$

Combining these relations and using the linearity property of the flux and triple integrals, we will arrive at the *Ostrogradsky formula*

$$\oiint\limits_{S} (P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}, \mathbf{n}^0)\, d\sigma = \iiint\limits_{V} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dv. \blacktriangleright$$

In closing, we should note that the Ostrogradsky formula also holds when the domain V is multiply connected.

Examples. (1) Find the flux of the vector $\mathbf{a} = 2x\mathbf{i} - (z - 1)\mathbf{k}$ through the closed surface S: $\{x^2 + y^2 = 4, z = 0, z = 1\}$ (1) by definition, (2) by the Ostrogradsky formula.

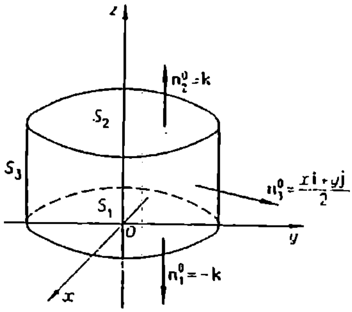


Fig. 24.24

◄ (1) By definition the flux of $\mathbf{a}$ is (Fig. 24.24)

$$\Pi = \Pi_1 + \Pi_2 + \Pi_3,$$

where

$$\Pi_1 = \iint_{S_1} (\mathbf{a}, \mathbf{n}_1^0)\, d\sigma = \iint_{S_1} (z - 1)\, d\sigma = -\iint_{S_1} d\sigma = -4\pi$$

since on S_1 we have $z = 0$;

$$\Pi_2 = \iint_{S_2} (\mathbf{a}, \mathbf{n}_2^0)\, d\sigma = -\iint_{S_2} (z - 1)\, d\sigma = 0,$$

since on S_2 we have $z = 1$;

$$\Pi_3 = \iint_{S_3} (\mathbf{a}, \mathbf{n}_3^0)\, d\sigma = \iint_{S_3} x^2\, d\sigma,$$

since

$$\mathbf{n}_3^0 = \frac{\operatorname{grad}(x^2 + y^2 - 4)}{|\operatorname{grad}(x^2 + y^2 - 4)|} = \frac{x\mathbf{i} + y\mathbf{j}}{2}.$$

We change to curvilinear coordinates on the cylinder

$$x = 2\cos\varphi, \quad y = 2\sin\varphi, \quad z = z, \quad d\sigma = 2\, d\varphi\, dz.$$

Then

$$\Pi_3 = \int_0^{2\pi} d\varphi \int_0^1 4\cos^2\varphi \cdot 2\, dz = 4\int_0^{2\pi} (1 + \cos 2\varphi)\, d\varphi = 8\pi.$$

Hence

$$\Pi = -4\pi + 0 + 8\pi = 4\pi.$$

(2) By the Ostrogradsky formula

$$\Pi = \iiint\limits_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dv = \iiint\limits_V (2 + 0 - 1)\, dv$$

$$= \iiint\limits_V dv = 4\pi. \blacktriangleright$$

(2) Find the flux of the radius vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ through the sphere of radius R and centre at the origin of coordinates (1) by definition and (2) by the Ostrogradsky formula.

◄ (1) Since for the sphere we have $\mathbf{n}^0 = (x\mathbf{i} + y\mathbf{j} + z\mathbf{k})/R$, then $(\mathbf{r}, \mathbf{n}^0) = (x^2 + y^2 + z^2)/R = R$, and therefore

$$\Pi = \oiint\limits_{S_R} (\mathbf{r}, \mathbf{n}^0)\, d\sigma = R \iint\limits_{S_R} d\sigma = 4\pi R^3.$$

(2) By the Ostrogradsky formula we find

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = 3,$$

then

$$\Pi = \iiint\limits_V 3\, dv = 3 \cdot \frac{4}{3} \pi R^3 = 4\pi R^3. \blacktriangleright$$

(3) Find the flux of the vector $\mathbf{a} = 3x\mathbf{i} - y\mathbf{j} - z\mathbf{k}$ through the closed surface S: $\begin{cases} 9 - z = x^2 + y^2 \\ z = 0 \quad (z > 0) \end{cases}$ (Fig. 24.25) (1) by definition and (2) by the Ostrogradsky formula.

◄ (1) By definition the flux of $\mathbf{a}$ will be

$$\Pi = \Pi_1 + \Pi_2,$$

where

$$\Pi_1 = \iint\limits_{S_1} (\mathbf{a}, \mathbf{n}_1^0)\, d\sigma = \iint\limits_{S_1} z\, d\sigma,$$

since on S_1 we have $z = 0$,

$$\Pi_2 = \iint\limits_{D_{xy}} (\mathbf{a}, \mathbf{n}_2^0)\Big|_{z = f(x,y)} dx\, dy,$$

$$\mathbf{n}_2^0 = \operatorname{grad}(x^2 + y^2 + z - 9) = 2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k},$$

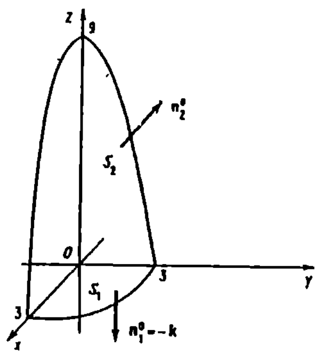


Fig. 24.25

and D_{xy} is the circle $x^2 + y^2 \leqslant 9$. Hence

$$(\mathbf{a}, \mathbf{n}_2^0)\Big|_{z=9-x^2-y^2} = (6x^2 - 2y^2 - z)\Big|_{z=9-x^2-y^2} = 7x^2 - y^2 - 9.$$

Therefore,

$$\Pi_2 = \iiint\limits_{D_{xy}} (7x^2 - y^2 - 9)\,dx\,dy$$

$$= \int_0^{2\pi} d\varphi \int_0^3 (7\varrho^2\cos^2\varphi - \varrho^2\sin^2\varphi - 9)\varrho\,d\varrho$$

$$= \int_0^{2\pi} \left(8\,\frac{\varrho^4}{4}\cos^2\varphi - \frac{\varrho^4}{4} - 9\,\frac{\varrho^2}{2}\right)\Bigg|_{\varrho=0}^{3} d\varphi$$

$$= \int_0^{2\pi} \left(162\cos^2\varphi - \frac{81}{4} - \frac{81}{2}\right) d\varphi$$

$$= \int_0^{2\pi} \left(81 + 81\cos 2\varphi - \frac{243}{4}\right) d\varphi = \frac{81}{4}\,2\pi = \frac{81}{4}\,\pi,$$

since $\int\limits_0^{2\pi} \cos 2\varphi\,d\varphi = 0$.

Thus $\Pi = \Pi_1 + \Pi_2 = \frac{81}{2}\pi$.

(2) By the Ostrogradsky formula we have

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = 3 - 1 - 1 = 1.$$

Therefore

$$\Pi = \iiint\limits_V 1\, dv.$$

Passing over to cylindrical coordinates

$$x = \varrho\cos\varphi, \qquad y = \varrho\sin\varphi, \qquad z = z, \qquad dv = \varrho\, d\varrho\, d\varphi\, dz$$

we obtain $z = 9 - \varrho^2$ from the equation of S. Hence

$$\Pi = \int\limits_0^{2\pi} d\varphi \int\limits_0^3 \varrho\, d\varrho \int\limits_0^{9-\varrho^2} dz = 2\pi \int\limits_0^3 (9 - \varrho^2)\varrho\, d\varrho$$

$$= 2\pi\left(9\frac{\varrho^2}{2} - \frac{\varrho^4}{4}\right)\Bigg|_0^3 = \frac{81}{2}\pi. \blacktriangleright$$

Remark. When dealing with an open surface S it is often convenient to close the surface and to make use of the Ostrogradsky-Gauss formula.

(4) Find the flux of the vector $\mathbf{a} = (y^2 + z^2)\mathbf{i} - y^2\mathbf{j} + 2yz\mathbf{k}$ through the surface S: $x^2 + z^2 = y^2$ $(0 \leqslant y \leqslant 1)$.

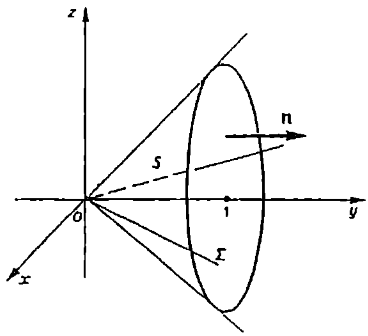


Fig. 24.26

◀ Surface S is the surface of a cone with the y-axis as its axis (Fig. 24.26). We close the cone by a piece Σ of the plane $y = 1$. If then we denote by Π_1 the flux we seek, by Π_2 the flux through Σ, we will have

$$\Pi_1 + \Pi_2 = \iiint\limits_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dv,$$

where V is the volume of the cone bounded by S and Σ. Since

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = 0 - 2y + 2y = 0,$$

we have $\Pi_1 + \Pi_2 = 0$, i.e., $\Pi_1 = -\Pi_2$, where

$$\Pi_2 = \iint_{\Sigma} (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iint_{\Sigma} (\mathbf{a}, \mathbf{j})\, d\sigma = \iint_{\Sigma} (-y^2)\, d\sigma = -\iint_{\Sigma} d\sigma = -\pi,$$

because on Σ we have $y = 1$. Hence $\Pi_1 = \pi$.

24.7 Divergence of a Vector Field

Solenoidal fields. To be more specific, we will look at the velocity field $\mathbf{v}$ of the flow of a fluid. Let S be a closed surface. If the flux through S is positive, this suggests that for the space bounded by S the fluid output is larger than the input. It is then said that there are *sources* within S that generate the fluid.

Conversely, if the flow is negative, the input is larger than the output. We say now that there are *sinks* within S that absorb the fluid.

Consequently, the quantity

$$\Pi = \oiint_{S} (\mathbf{a}, \mathbf{n}^0)\, d\sigma$$

tells us about the nature of the vector field confined within the surface S, namely about the presence of sources or sinks within the space and about their contributions. The concept of the flux of a vector through a closed surface leads to the concept of the *divergence of a field*. It is a quantitative characteristic of the field at each point of space.

Let M be a given point of the field. We surround it with a surface S of arbitrary shape, e.g., by the sphere of a sufficiently small radius. We will denote by (V) the region bounded by S and its volume by V.

Let us determine the flux of vector $\mathbf{a}$ through S. We will have

$$\Pi = \oiint_{S} (\mathbf{a}, \mathbf{n}^0)\, d\sigma.$$

Consider the ratio

$$\frac{\oiint_{S} (\mathbf{a}, \mathbf{n}^0)\, d\sigma}{V}. \tag{24.67}$$

Since the numerator is the input of the sources inside (V), then the ratio (24.67) yields the mean input of a unit volume.

Definition. If (24.67) has a finite limit when (V) shrinks to point M, then this limit is called the *divergence of the vector field* (divergence of a) at M and is denoted by $\operatorname{div} \mathbf{a}(M)$. And so by definition

$$\operatorname{div} \mathbf{a}(M) = \lim_{(V)\to M} \frac{\oint\!\!\oint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma}{V}. \tag{24.68}$$

The divergence of a vector field is a scalar quantity, since both the numerator and denominator in (24.68) are scalar quantities.

If $\operatorname{div} \mathbf{a}(M) > 0$, then at the point M we have a source, if $\operatorname{div} \mathbf{a}(M) < 0$, then at the point we have a sink.

From formula (24.67), which defines the divergence, we conclude that the divergence of the field of $\mathbf{a}$ at the point M is the volume density of the flux of $\mathbf{a}$ at that point.

Formula (24.68) is the invariant definition of the divergence, i.e., a definition unaffected by the choice of the system of coordinates, since all the quantities in (24.68) are determined directly by the field itself and are independent of the choice of the coordinate system.

We show next how to compute the divergence in rectangular coordinates provided that the vector coordinates

$$\mathbf{a} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

are continuous and have continuous partial derivatives $\partial P/\partial x$, $\partial Q/\partial y$, $\partial R/\partial z$ in a neighbourhood of the point M.

Under these conditions we can apply the Ostrogradsky-Gauss theorem to the flux of $\mathbf{a}$ through any closed surface S in the vicinity of M

$$\oint\!\!\oint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iiint_{(V)} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dv.$$

To the triple integral on the right we will apply the mean value theorem

$$\oint\!\!\oint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)_{M_m} V.$$

Substituting this into (24.68), we will find

$$\operatorname{div} \mathbf{a}(M) = \lim_{(V)\to M} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)_{M_m}.$$

When the region (V) shrinks to a point M, the point M_m too tends to M, and on the assumption that the partial derivatives are continuous we

obtain

$$\operatorname{div} \mathbf{a}(M) = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)_M$$

or, for short,

$$\operatorname{div} \mathbf{a} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}. \tag{24.69}$$

We assume that all the quantities in (24.69) are taken at the same point. Expression (24.69) yields the divergence in rectangular coordinates. We have proved by the way that the divergence of $\mathbf{a}$ exists provided that $\partial P/\partial x$, $\partial Q/\partial y$, and $\partial R/\partial z$ are continuous. Using (24.69), we will write the Ostrogradsky formula, in vector, form

$$\oint\!\!\oint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iiint_V \operatorname{div} \mathbf{a}\, dv, \tag{24.70}$$

i.e., the flux of $\mathbf{a}$ through a closed surface S is equal to the triple integral of the divergence of $\mathbf{a}$ over (V) bounded by S.

Divergence is computed using the following rules:

(1) The divergence has the *linearity* property

$$\operatorname{div}(C_1\mathbf{a}_1 + \ldots + C_n\mathbf{a}_n) = C_1 \operatorname{div} \mathbf{a}_1 + \ldots + C_n \operatorname{div} \mathbf{a}_n, \tag{24.71}$$

where $C_1, \ldots, C_n$ are numerical constants.

◄ Let $\mathbf{a} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$ and C be a constant. Then

$$\operatorname{div} C_{\mathbf{a}} = \operatorname{div}(CP\mathbf{i} + CQ\mathbf{j} + CR\mathbf{k}) = C\frac{\partial P}{\partial x} + C\frac{\partial Q}{\partial y} + C\frac{\partial R}{\partial z} = C \operatorname{div} \mathbf{a}.$$

Further, if

$$\mathbf{a}_1 = P_1\mathbf{i} + Q_1\mathbf{j} + R_1\mathbf{k}, \qquad \mathbf{a}_2 = P_2\mathbf{i} + Q_2\mathbf{j} + R_2\mathbf{k},$$

then

$$\begin{aligned}\operatorname{div}(\mathbf{a}_1 + \mathbf{a}_2) &= \operatorname{div}[(P_1 + P_2)\mathbf{i} + (Q_1 + Q_2)\mathbf{j} + (R_1 + R_2)\mathbf{k}] \\ &= \left(\frac{\partial P_1}{\partial x} + \frac{\partial P_2}{\partial x}\right) + \left(\frac{\partial Q_1}{\partial y} + \frac{\partial Q_2}{\partial y}\right) + \left(\frac{\partial R_1}{\partial z} + \frac{\partial R_2}{\partial z}\right) \\ &= \left(\frac{\partial P_1}{\partial x} + \frac{\partial Q_1}{\partial y} + \frac{\partial R_1}{\partial z}\right) + \left(\frac{\partial P_2}{\partial x} + \frac{\partial Q_2}{\partial y} + \frac{\partial R_2}{\partial z}\right) \\ &= \operatorname{div} \mathbf{a}_1 + \operatorname{div} \mathbf{a}_2 \blacktriangleright\end{aligned}$$

(2) The divergence of a constant vector $\mathbf{c}$ is zero

$$\operatorname{div} \mathbf{c} = 0, \tag{24.72}$$

since all the coordinates of $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$ are constant.

(3) The divergence of the product of a scalar function $u(M)$ and a vector $\mathbf{a}(M)$ is found by the formula

$$\operatorname{div}(u\mathbf{a}) = u \operatorname{div} \mathbf{a} + (\mathbf{grad}\, u, \mathbf{a}). \qquad (24.73)$$

◄ Really

$$\begin{aligned}\operatorname{div}(u\mathbf{a}) &= \operatorname{div}[uP\mathbf{i} + uQ\mathbf{j} + uR\mathbf{k}] \\ &= \frac{\partial(uP)}{\partial x} + \frac{\partial(uQ)}{\partial y} + \frac{\partial(uR)}{\partial z} \\ &= u\frac{\partial P}{\partial x} + u\frac{\partial Q}{\partial y} + u\frac{\partial R}{\partial z} + \frac{\partial u}{\partial x}P + \frac{\partial u}{\partial y}Q + \frac{\partial u}{\partial z}R \\ &= u \operatorname{div} \mathbf{a} + (\mathbf{grad}\, u, \mathbf{a}). ►\end{aligned}$$

Example. Find the divergence of

$$\mathbf{a} = \varphi(r)\mathbf{r}^0 = \varphi(r)\frac{\mathbf{r}}{r},$$

where $r = |\mathbf{r}|$ is the distance from the origin to a variable point $M(x, y, z)$, $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.
◄ By (24.73)

$$\operatorname{div} \mathbf{a} = \frac{\varphi(r)}{r}\operatorname{div}\mathbf{r} + \left(\mathbf{grad}\frac{\varphi(r)}{r}, \mathbf{r}\right).$$

Since $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, we have

$$\operatorname{div}\mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3.$$

Further,

$$\begin{aligned}\mathbf{grad}\frac{\varphi(r)}{r} &= \left(\frac{\varphi(r)}{r}\right)'_r \mathbf{grad}\, r = \frac{r\varphi'(r) - \varphi(r)}{r^2}\left(\frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{\sqrt{x^2 + y^2 + z^2}}\right) \\ &= \frac{r\varphi'(r) - \varphi(r)}{r^2}\mathbf{r}^0,\end{aligned}$$

therefore,

$$\begin{aligned}\left(\mathbf{grad}\frac{\varphi(r)}{r}, \mathbf{r}\right) &= \left(\frac{r\varphi'(r) - \varphi(r)}{r^2}\mathbf{r}^0, \mathbf{r}\right) \\ &= \frac{r\varphi'(r) - \varphi(r)}{r} = \varphi'(r) - \frac{\varphi(r)}{r}.\end{aligned}$$

Thus

$$\operatorname{div}\mathbf{a} = 3\frac{\varphi(r)}{r} + \varphi'(r) - \frac{\varphi(r)}{r} = 2\frac{\varphi(r)}{r} + \varphi'(r). ►$$

If at all points of some domain G the divergence of a vector field defined in G is zero, i.e.,

$$\operatorname{div} \mathbf{a} \equiv 0 \tag{24.74}$$

in the domain G, then the field is said to be *solenoidal* in that domain.

It follows from the Ostrogradsky-Gauss theorem that in a solenoidal field the flux of a vector through any closed surface S contained in that field is zero

$$\oint\!\!\oint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = 0. \tag{24.75}$$

In fact, by the Ostrogradsky formula

$$\oint\!\!\oint_S (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iiint_V \operatorname{div} \mathbf{a}\, dv$$

and since we have assumed that $\operatorname{div} \mathbf{a} = 0$, (24.75) is valid.

Properties of a solenoidal field. Consider some plane region Σ in the field of a vector $\mathbf{a}$. The totality of the vector lines passing through the boundary γ of that area is called the *vector tube*. Let Σ_1 be some cross-section of the tube. The normal $\mathbf{n}_1$ to Σ_1 is oriented in the same direction as the vector $\mathbf{a}$ of the field (Fig. 24.27).

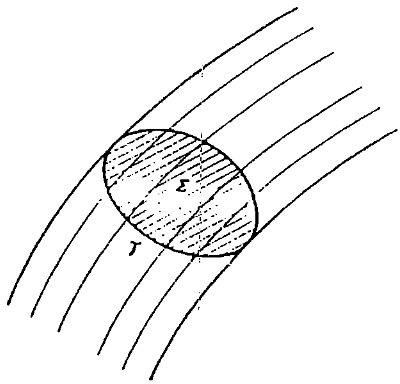


Fig. 24.27

Theorem 24.5. *In a solenoidal field the flux of vector* $\mathbf{a}$ *through any section of the vector tube is the same.*

◄ Let Σ_1 and Σ_2 be two nonintersecting sections of the same vector tube. We have to prove that

$$\iint_{\Sigma_1} (\mathbf{a}, \mathbf{n}_1)\, d\sigma = \iint_{\Sigma_2} (\mathbf{a}, \mathbf{n}_2)\, d\sigma.$$

We denote by Σ_3 the part of the surface of the tube contained between the sections Σ_1 and Σ_2. The surfaces Σ_1, Σ_2, Σ_3 together form the closed surface Σ (Fig. 24.28).

Since the field of **a** is assumed to be solenoidal, we have

$$\oiint_{\Sigma} (\mathbf{a}, \mathbf{n}^0)\, d\sigma = 0. \tag{24.76}$$

The flux being additive, we rewrite (24.76) as

$$\iint_{\Sigma_1} (\mathbf{a}, \mathbf{n}_1^0)\, d\sigma + \iint_{\Sigma_2} (\mathbf{a}, -\mathbf{n}_2)\, d\sigma + \iint_{\Sigma_3} (\mathbf{a}, \mathbf{n}_3^0)\, d\sigma = 0. \tag{24.77}$$

On the surface Σ_3, which is made up of vector lines, we have $\mathbf{n}_3^0 \perp \mathbf{a}$, so that $(\mathbf{a}, \mathbf{n}_3^0) = 0$ in Σ_3, and hence the last integral on the left-hand side of (24.77) is zero.

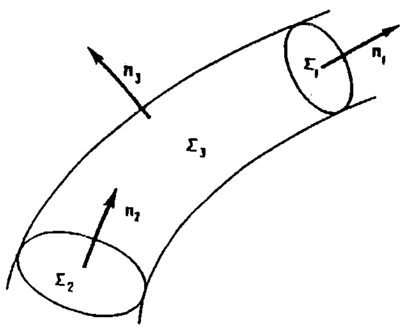


Fig. 24.28

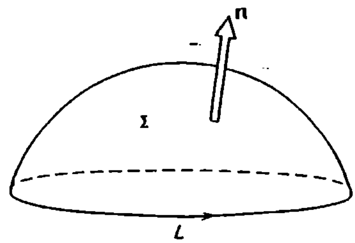


Fig. 24.29

From (24.77) we thus find

$$\iint_{\Sigma_1} (\mathbf{a}, \mathbf{n}_1^0)\, d\sigma = \iint_{\Sigma_2} (\mathbf{a}, \mathbf{n}_2^0)\, d\sigma. \blacktriangleright$$

Let L be an oriented closed contour and let it be the boundary of Σ. We will say that the surface Σ is stretched on the contour L. We will orient the normal **n** to Σ so that if looking from the end of the normal the contour would be traced counterclockwise. In other words, if we move over the surface round the contour so that the normal chosen would point upwards, then, as we trace the contour in the direction chosen, the nearest part of Σ will be on the left (Fig. 24.29).

Theorem 24.6. *In a solenoidal field the flux of vector* **a** *through any surface stretched on a given contour is the same*

$$\iint_{\Sigma_1} (\mathbf{a}, \mathbf{n}_1^0)\, ds = \iint_{\Sigma_2} (\mathbf{a}, \mathbf{n}_2^0)\, ds. \tag{24.78}$$

We leave it for the reader to prove formula (24.78) (Fig. 24.30).

Remark. In a solenoidal field vector lines can neither originate nor terminate within the field. They can be either closed curves or have their ends on the boundary of the domain where the field is defined.

Example. Consider a field of force produced by a point charge q placed at the origin of coordinates. The strength of the field is

$$\mathbf{E} = \frac{q}{r^2}\mathbf{r}^0. \tag{24.79}$$

◄ We have

$$\operatorname{div}\mathbf{E} = \operatorname{div}\frac{q}{r^2}\mathbf{r}_0 = \operatorname{div}\left(\frac{q}{r^3}\mathbf{r}\right),$$

where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, $r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$.

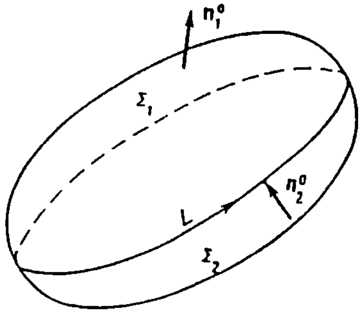


Fig. 24.30

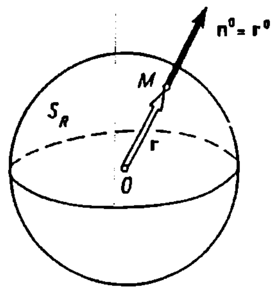


Fig. 24.31

Further, by (24.73) we have (for $r \neq 0$)

$$\operatorname{div}\left(\frac{q}{r^3}\mathbf{r}\right) = \frac{q}{r^3}\operatorname{div}\mathbf{r} + \left(\operatorname{grad}\frac{q}{r^3}, \mathbf{r}\right) = \frac{3q}{r^3} + \left(-\frac{3q}{r^4}\mathbf{r}^0, r\mathbf{r}^0\right)$$
$$= \frac{3q}{r^3} - \frac{3q}{r^3} = 0.$$

The field of $\mathbf{E}$ given by (24.79) will be solenoidal in any domain G that does not contain the point $O(0, 0, 0)$.

Consider the flux of $\mathbf{E}$ through the sphere S_R of radius R and centre at the origin of coordinates $O(0, 0, 0)$ (Fig. 24.31)

$$\Pi = \oint\!\!\oint_{S_R} (\mathbf{E}, \mathbf{n}^0)\,d\sigma = \iint_{S_R}\left(\frac{q}{r^2}\mathbf{r}^0, \mathbf{r}^0\right)d\sigma = \frac{q}{R^2}\iint_{S_R} d\sigma = \frac{q}{R^2}4\pi R^2 = 4\pi q,$$

since on S_R we have $\mathbf{E} = q\mathbf{r}^0/R^2$ and $\oint\!\!\oint_{S_R} d\sigma = 4\pi R^2$. ►

It can be shown that the flux of $\mathbf{E}$ through any closed surface that contains $O(0, 0, 0)$ is $4\pi q$.

24.8 Circulation of a Vector Field. Curl of a Vector. Stokes Theorem

Suppose that in some domain G we have a continuous vector field

$$\mathbf{a}(M) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

and a closed oriented contour L.

Definition. The *circulation* of a vector $\mathbf{a}$ over a closed contour L is the line integral of the second kind of $\mathbf{a}$ along L, i.e.,

$$\text{circulation} = \oint_L (\mathbf{a}, d\mathbf{r}) = \oint_L (P\,dx + Q\,dy + R\,dz). \qquad (24.80)$$

Here $d\mathbf{r}$ is a vector whose magnitude equals the differential of the arc L and direction coincides with that of the tangent to L, which in turn depends on the orientation of the contour (Fig. 24.32).

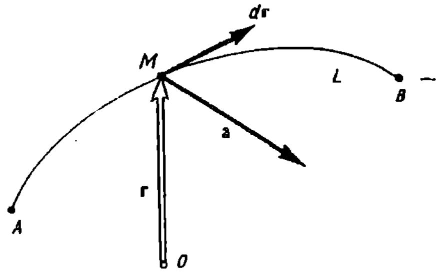


Fig. 23.32

The symbol $\oint_L$ stands for the integral around the closed contour L.

Example. Find the circulation of the vector field $\mathbf{a} = -y^3\mathbf{i} + x^3\mathbf{j}$ along the ellipse L: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$.

◀ By the definition of circulation we have

$$\text{circulation} = \oint_L (\mathbf{a}, d\mathbf{r}) = \oint_L -y^3\,dx + x^3\,dy. \qquad (24.81)$$

The parametric equations of the ellipse have the form

$$\begin{cases} x = a\cos t \\ y = b\sin t \end{cases} \quad 0 \leqslant t \leqslant 2\pi,$$

and so $dx = -a\sin t\,dt$, $dy = b\cos t\,dt$.

Substituting this into (24.81) gives

$$\text{circulation} = ab \int_0^{2\pi} (b^2 \sin^4 t + a^2 \cos^4 t) = \frac{3\pi ab}{4} (a^2 + b^2),$$

since

$$\int_0^{2\pi} \sin^4 t\, dt = \frac{1}{4} \int_0^{2\pi} (1 - \cos 2t)^2\, dt$$
$$= \frac{1}{4} \int_0^{2\pi} \left(1 - 2\cos 2t + \frac{1 + \cos 4t}{2}\right) dt$$
$$= \frac{1}{4} \int_0^{2\pi} \left(\frac{3}{2} - 2\cos 2t + \frac{1}{2}\cos 4t\right) dt = \frac{3}{4}\pi.$$

Similarly, we find that

$$\int_0^{2\pi} \cos^4 t\, dt = \frac{3}{4}\pi. \blacktriangleright$$

Curl of a vector field. Suppose that we have the field of a vector

$$\mathbf{a}(M) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}.$$

Suppose further that the coordinates P, Q, R of $\mathbf{a}(M)$ are continuous and have continuous partial derivatives of the first order in all the arguments.

Definition. The *curl* of a vector $\mathbf{a}(M)$ is the vector denoted by **curl a** and given by

$$\mathbf{curl\, a} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k},$$

or, in shorthand notation,

$$\mathbf{curl\, a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}. \tag{24.82}$$

This determinant is expanded in the elements of the first row; the multiplication of the elements of the second row by the elements of the third row

is understood as differentiation, e.g., $\frac{\partial}{\partial x} Q = \frac{\partial Q}{\partial x}$, $\mathbf{i}\begin{vmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ Q & R \end{vmatrix} =$ $\mathbf{i}\left(\frac{\partial}{\partial y}R - \frac{\partial}{\partial z}Q\right) = \mathbf{i}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)$.

Definition. If in some domain G we have $\mathbf{curl\,a} = 0$, then the field of $\mathbf{a}$ in G is called *irrotational*.

Example. Find the curl of $\mathbf{a} = -\frac{y^2}{2}\mathbf{i} + \frac{x^2}{2}\mathbf{j}$.

◄ According to (24.82) we have

$$\mathbf{curl\,a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y^2/2 & x^2/2 & 0 \end{vmatrix} = 0 \cdot \mathbf{i} + 0 \cdot \mathbf{j} + (x + y)\mathbf{k}. \text{►}$$

Since $\mathbf{curl\,a}$ is a vector, we can also consider its vector field, i.e., the field of the curl of $\mathbf{a}$.

Assuming that the coordinates of $\mathbf{a}$ have continuous partial derivatives of the second order, we compute the divergence of $\mathbf{curl\,a}$. The result will be

$$\begin{aligned} \operatorname{div}\mathbf{curl\,a} &= \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \\ &= \frac{\partial^2 R}{\partial y\,\partial x} - \frac{\partial^2 Q}{\partial z\,\partial x} + \frac{\partial^2 P}{\partial z\,\partial y} - \frac{\partial^2 R}{\partial x\,\partial y} + \frac{\partial^2 Q}{\partial x\,\partial z} - \frac{\partial^2 P}{\partial y\,\partial z} = 0 \end{aligned}$$

i.e., $\operatorname{div}\mathbf{curl\,a} = 0$. (24.82′)

The field of $\mathbf{curl\,a}$ is thus a *solenoidal* field.

Theorem 24.7 (Stokes' theorem). *The circulation of vector* **a** *around an oriented closed contour L equals the flux of the curl of the vector through any surface strectched onto L:*

$$\oint_L (\mathbf{a}, d\mathbf{r}) = \iint_\Sigma (\mathbf{curl\,a}, \mathbf{n}^0)\, d\sigma. \tag{24.83}$$

It is supposed here that the coordinates of $\mathbf{a}$ have continuous partial derivatives in some space domain G and that the orientation of the normal unit vector $\mathbf{n}^0$ to $\Sigma \subset G$ is consistent with the orientation of L so that from the end of the normal the contour appears to be traced counterclockwise.

◄ Considering that $\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ and $\mathbf{n}^0 = \mathbf{i}\cos\alpha + \mathbf{j}\cos\beta + \mathbf{k}\cos\gamma$ and using the definition (24.82) of the curl, we rewrite (24.83) as follows

$$\oint_L P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz$$

$$= \iint_{\Sigma} \left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \cos\alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \cos\beta + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \cos\gamma \right] d\sigma$$

$$= \iint_{\Sigma} \left[\left(\frac{\partial P}{\partial z} \cos\beta - \frac{\partial P}{\partial y} \cos\gamma \right) + \left(\frac{\partial Q}{\partial x} \cos\gamma - \frac{\partial Q}{\partial z} \cos\alpha \right) + \left(\frac{\partial R}{\partial y} \cos\alpha - \frac{\partial R}{\partial x} \cos\beta \right) \right] d\sigma. \quad (24.84)$$

Consider first the case where the smooth surface Σ and its contour L are uniquely projectable on the xy-plane in a domain D and its boundary — contour λ, respectively. The orientation of the contour L causes the contour λ to be oriented in a definite manner.

For definiteness, we will consider that L is oriented so that the surface Σ lies to the left. The normal $\mathbf{n}$ to Σ will then make an acute angle γ with the z-axis ($\cos\gamma > 0$).

Let Σ be desribed by the equation $z = \varphi(x, y)$, where the function $\varphi(x, y)$ is continuous and has continuous partial derivatives $\partial\varphi/\partial x$ and $\partial\varphi/\partial y$ in the closed domain D.

Consider the integral

$$\oint_L P(x, y, z)\, dx. \quad (24.85)$$

The line L lies on the surface Σ and, using the equation of the surface $z = \varphi(x, y)$, we can substitute $\varphi(x, y)$ for z in (24.85). The integrand $P(x, y, \varphi(x, y))$ will then contain only x and y. The coordinates (x, y) of the variable point of the curve λ are the same as those of the corresponding point of the curve L, and so integration along L can be replaced by integration along λ (Fig. 24.33)

$$\oint_L P(x, y, z)\, dx = \oint_\lambda P(x, y, \varphi(x, y))\, dx. \quad (24.86)$$

We apply to the integral on the right Green's theorem

$$\oint_\lambda P(x, y, \varphi(x, y))\, dx = -\iint_D \left[\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} \frac{\partial \varphi}{\partial y} \right] ds. \quad (24.87)$$

We now pass over from the integral over the space domain D to the integral over the surface Σ. Since $ds = \cos\gamma\, d\sigma$, then from (24.87) we obtain

$$\oint_L P(x, y, z)\, dx = -\iint_\Sigma \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} \frac{\partial \varphi}{\partial y} \right) \cos\gamma\, d\sigma. \quad (24.88)$$

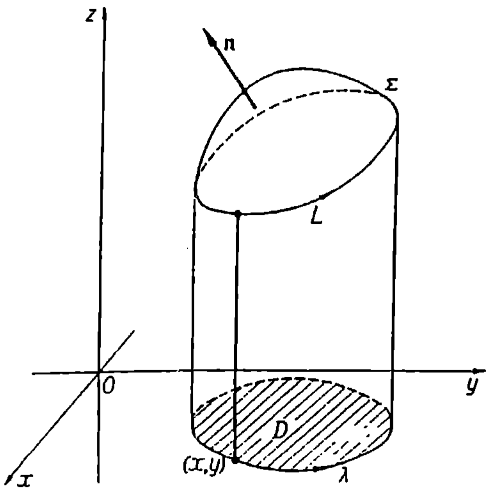


Fig. 24.33

The normal $\mathbf{n}^0$ to Σ is given by

$$\mathbf{n}^0 = \frac{\mathbf{grad}\,(z - \varphi(x, y))}{|\mathbf{grad}\,(z - \varphi(x, y))|} = \frac{-\dfrac{\partial\varphi}{\partial x}\mathbf{i} - \dfrac{\partial\varphi}{\partial y}\mathbf{j} + \mathbf{k}}{\sqrt{\left(\dfrac{\partial\varphi}{\partial x}\right)^2 + \left(\dfrac{\partial\varphi}{\partial y}\right)^2 + 1}},$$

or

$$\mathbf{n}^0 = \mathbf{i}\cos\alpha + \mathbf{j}\cos\beta + \mathbf{k}\cos\gamma.$$

We notice that $(\partial\varphi/\partial y)\cos\gamma = -\cos\beta$. Therefore, we can rewrite (24.88) as

$$\oint_L P(x, y, z)\,dx = -\iint_\Sigma \left(\frac{\partial P}{\partial y}\cos\gamma - \frac{\partial P}{\partial z}\cos\beta\right) d\sigma$$
$$= \iint_\Sigma \left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right) d\sigma. \qquad (24.89)$$

Likewise, if we assume that Σ is a smooth surface that is uniquely projectable on all the three coordinate planes, we see that the following formulas hold:

$$\oint_L Q(x, y, z)\,dy = \iint_\Sigma \left(\frac{\partial Q}{\partial x}\cos\gamma - \frac{\partial Q}{\partial z}\cos\alpha\right) d\sigma, \qquad (24.90)$$

$$\oint_L R(x, y, z)\,dz = \iint_\Sigma \left(\frac{\partial R}{\partial y}\cos\alpha - \frac{\partial R}{\partial x}\cos\beta\right) d\sigma. \qquad (24.91)$$

Adding up (24.89), (24.90) and (24.91) term by term, we obtain Stokes' formula (24.84), or for short

$$\oint_L (\mathbf{a},\, d\mathbf{r}) = \iint_\Sigma (\operatorname{curl} \mathbf{a},\, \mathbf{n}^0)\, d\sigma. \blacktriangleright$$

Remarks. (1) We have shown that the field of **curl a** is a solenoidal field, and so the flux of **curl a** is independent of the kind of surface Σ stretched on the contour L.

(2) Formula (24.83) is derived on the assumption that Σ is uniquely projectable on all the three coordinate planes. If this condition is not met, we break up Σ into parts so that each part would meet the condition. And then we use the additivity property of integrals.

Example. Compute the circulation of the vector $\mathbf{a} = y\mathbf{i} - x\mathbf{j} + \mathbf{k}$ along the line L: $\begin{cases} x^2 + y^2 = R^2 \\ z = H \end{cases}$ $(H > 0)$ using (1) the definition and (2) Stokes' theorem.

◄ (1) We will define L parametrically, i.e., L: $\{x = R\cos t,\ y = R\sin t,\ z = H,\ 0 \leqslant t \leqslant 2\pi\}$. Then $dx = -R\sin t\, dt$, $dy = R\cos t\, dt$, $dz = 0$, so that the circulation will be

$$\oint_L y\,dx - x\,dy + dz = \int_0^{2\pi} (-R^2\sin^2 t - R^2\cos^2 t)\,dt = -2\pi R^2.$$

(2) Find **curl a**.

$$\operatorname{curl}\mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ y & -x & 1 \end{vmatrix} = 0\cdot\mathbf{i} + 0\cdot\mathbf{j} - 2\mathbf{k} = -2\mathbf{k}.$$

We now stretch the contour L on the plane $z = H$ so that $\mathbf{n}^0 = \mathbf{k}$. Then the circulation will be

$$\iint_\Sigma (\operatorname{curl}\mathbf{a},\, \mathbf{n}^0)\, d\sigma = -2\iint_\Sigma d\sigma = -2\pi R^2.$$

Invariant definition of the curl of a field. From the Stokes' theorem we can obtain the definition independent of the choice of a coordinate system.

Theorem 24.8. *The projection of* **curl a** *on any direction is independent of the choice of a coordinate system and is equal to the surface density of the circulation of* **a** *around the contour of the area perpendicular to this direction*,

$$\operatorname{pr}_{\mathbf{n}} \operatorname{curl}\mathbf{a}\,|_M = (\operatorname{curl}\mathbf{a},\, \mathbf{n}^0)\,|_M = \lim_{(\Sigma)\to M} \frac{\oint_L (\mathbf{a},\, d\mathbf{r})}{S}. \qquad (24.92)$$

Here (Σ) is a plane patch perpendicular to n, S is the area of the patch, L is the contour of the patch oriented so that when seen from the end of **n** the contour would be traced counterclockwise; $(\Sigma) \to M$ means that the patch (Σ) shrinks to a point M where we consider **curl a**, the vector n on the patch always remaining the same (Fig. 24.34).

◄ We will first apply to the circulation

$$\oint_L (\mathbf{a}, d\mathbf{r})$$

of the vector **a** Stokes' theorem and then apply to the resultant double integral the mean value theorem to obtain

$$\oint_L (\mathbf{a}, d\mathbf{r}) = \iint_\Sigma (\mathbf{curl\,a}, \mathbf{n}^0)\, d\sigma = (\mathbf{curl\,a}, \mathbf{n}^0)\,|_{M_m} \cdot S.$$

Hence

$$\frac{\oint_L (\mathbf{a}, d\mathbf{r})}{S} = (\mathbf{curl\,a}, \mathbf{n}^0)\,|_{M_m},$$

where the scalar product $(\mathbf{curl\,a}, \mathbf{n}^0)$ is taken at some mean value point M_m on the patch (Σ).

As the patch (Σ) shrinks to M the point M_m also tends to M. Since we assumed that the partial derivatives of the coordinates of a (and hence curl a) are continuous, we arrive at

$$\lim_{(\Sigma)\to M} \frac{\oint_L (\mathbf{a}, d\mathbf{r})}{S} = \lim_{(\Sigma)\to M} (\mathbf{curl\,a}, \mathbf{n}^0)\,|_{M_m} = (\mathbf{curl\,a}, \mathbf{n}^0)\,|_M. \blacktriangleright$$

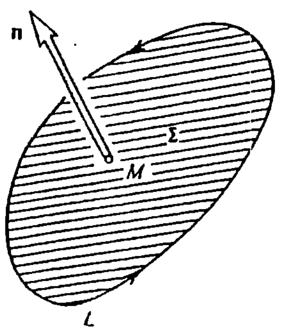


Fig. 24.34

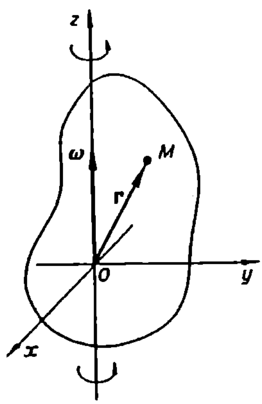


Fig. 24.35

Since the projection of **curl a** on an arbitrary direction is independent of the choice of a coordinate system, the vector **curl a** is invariant with respect of that choice. We thus have the following invariant definition of the curl of a field.

The *curl of a field* is a vector whose magnitude is equal to the largest surface density circulation at a given point; it is directed perpendicularly to the patch on which this largest density of circulation is achieved. The orientation of **curl a**, by the right-handed rule, is consistent with that of the contour on which the circulation is positive.

Physical meaning of the curl of a field. Consider a solid body rotating around its fixed axis l with an angular velocity ω. Without loss of generality we can assume the axis l to be coincident with the z-axis (Fig. 24.35).

Let $M(\mathbf{r})$ be the point under consideration, for which

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

The vector of the angular velocity is

$$\boldsymbol{\omega} = \omega\mathbf{k}.$$

The vector $\mathbf{v}$ of the linear velocity of a point M is

$$\mathbf{v} = [\boldsymbol{\omega}, \mathbf{r}] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 0 & \omega \\ x & y & z \end{vmatrix} = -y\omega\mathbf{i} + x\omega\mathbf{j}.$$

Hence

$$\text{curl}\,\mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y\omega & x\omega & 0 \end{vmatrix} = 2\omega\mathbf{k} = 2\boldsymbol{\omega}.$$

Thus

$$\text{curl}\,\mathbf{v} = 2\boldsymbol{\omega},$$

i.e., the curl of the field of velocities of a rotating solid is the same at all the points of the field, it is parallel to the axis of rotation and equal to twice the angular velocity of rotation.

Rules for computing the curl. (1) According to the definition (24.82) of the curl the latter is a vector. It is easily seen that the rotor of a constant vector $\mathbf{c}$ is zero

$$\text{curl}\,\mathbf{c} = \mathbf{0}.$$

(2) The curl has the *linearity* property

$$\operatorname{curl}(C_1\mathbf{a}_1 + C_2\mathbf{a}_2 + \ldots + C_n\mathbf{a}_n) = C_1 \operatorname{curl}\mathbf{a}_1 + C_2 \operatorname{curl}\mathbf{a}_2 + \ldots + C_n \operatorname{curl}\mathbf{a}_n,$$

where $C_1, C_2, \ldots, C_n$ are constant numbers.

This rule follows from the fact that the linearity property is inherent in all the derivatives that appear in the curl formula.

(3) The curl of the product of a scalar function $u(M)$ by a vector $\mathbf{a}(M)$ is computed by the formula

$$\operatorname{curl}(u\mathbf{a}) = u \operatorname{curl}\mathbf{a} + [\operatorname{grad} u,\ \mathbf{a}].$$

◄ We have

$$\operatorname{curl}(u\mathbf{a}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ uP & uQ & uR \end{vmatrix} = \left(\frac{\partial(uR)}{\partial y} - \frac{\partial(uQ)}{\partial z}\right)\mathbf{i}$$

$$+ \left(\frac{\partial(uP)}{\partial z} - \frac{\partial(uR)}{\partial x}\right)\mathbf{j} + \left(\frac{\partial(uQ)}{\partial x} - \frac{\partial(uP)}{\partial y}\right)\mathbf{k}$$

$$= u\left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}\right]$$

$$+ \left[\left(R\frac{\partial u}{\partial y} - Q\frac{\partial u}{\partial z}\right)\mathbf{i} + \left(P\frac{\partial u}{\partial z} - R\frac{\partial u}{\partial x}\right)\mathbf{j} + \left(Q\frac{\partial u}{\partial x} - P\frac{\partial u}{\partial y}\right)\mathbf{k}\right]$$

$$= u\operatorname{curl}\mathbf{a} + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ P & Q & R \end{vmatrix}$$

$$= u\operatorname{curl}\mathbf{a} + [\operatorname{grad} u,\ \mathbf{a}].$$ ►

24.9 Independence of the Line Integral of Integration Path

Definition. A domain G of three-dimensional space is said to be *simply connected on the surface* if on any closed contour lying in this domain we can stretch a surface that wholly belongs to G.

For example, the inside of the sphere, the whole of the three-dimensional space are simply connected domains, whereas the inside of the thorus, the three-dimensional space from which a straight line (e.g., the z-axis) is excluded are not simply connected domains.

Let G be simply connected on the surface. In G the continuous vector field $\mathbf{a}(M) = P(M)\mathbf{u} + Q(M)\mathbf{j} + R(M)\mathbf{k}$ is defined. Then the following theorem is valid:

Theorem 24.9. *The necessary and sufficient condition for the line integral*

$$\int_{AB} (\mathbf{a}, d\mathbf{r})$$

in the field of vector **a** *not to depend on the integration path, but only on the initial and terminal points (A and B) is that the circulation of* **a** *around any closed contour L lying in G be zero.*

◄ *Necessity.* Let $\int_{AB} (\mathbf{a}, d\mathbf{r})$ be independent of the integration path. We show that then $\oint_{L} (\mathbf{a}, d\mathbf{r})$ around any closed contour is zero.

Consider an arbitrary closed contour L in the field of **a** and take some arbitrary points A and B on it (Fig. 24.36).

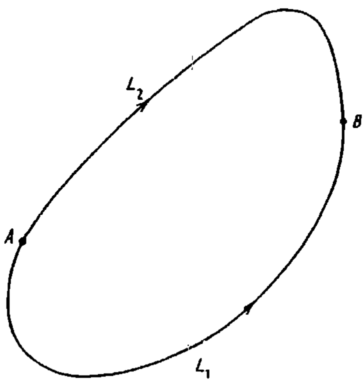


Fig. 24.36

As stated, we have

$$\int_{L_1} (\mathbf{a}, d\mathbf{r}) = \int_{L_2} (\mathbf{a}, d\mathbf{r}) = -\int_{L_2^-} (\mathbf{a}, d\mathbf{r}),$$

where L_1 and L_2 are different paths connecting points A and B. Hence

$$\int_{L_1 \cup L_2^-} (\mathbf{a}, d\mathbf{r}) = 0. \qquad (24.93)$$

But $L_1 \cup L_2^-$ is exactly the closed contour L we have chosen.

Sufficiency. Let $\oint_L (\mathbf{a}, d\mathbf{r})$ around any closed contour L be zero. Show that in that case $\int_{AB} (\mathbf{a}, d\mathbf{r})$ is independent of the integration path.

We will take two points A and B in the field of $\mathbf{a}$, connect them by arbitrary lines L_1 and L_2 and show that

$$\int_{L_1} (\mathbf{a}, d\mathbf{r}) = \int_{L_2} (\mathbf{a}, d\mathbf{r}). \tag{24.94}$$

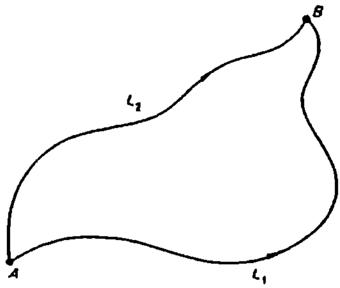


Fig. 24.37

We will only consider the simple case where the lines L_1 and L_2 do not meet. In that case $L_1 \cup L_2$ form a simple closed contour L (Fig. 24.37). We supposed that

$$\int_L (\mathbf{a}, d\mathbf{r}) = 0$$

but by the additivity property

$$\oint_L (\mathbf{a}, d\mathbf{r}) = \int_{L_1} (\mathbf{a}, d\mathbf{r}) + \int_{L_2^-} (\mathbf{a}, d\mathbf{r}),$$

and so

$$\int_{L_1} (\mathbf{a}, d\mathbf{r}) + \int_{L_2^-} (\mathbf{a}, d\mathbf{r}) = 0,$$

which proves the validity of (24.94). ►

The theorem expresses the necessary and sufficient condition for a line integral to be independent of the form of the path, but these conditions are hard to check. And so we will provide a more effective criterion.

Theorem 24.10. *For the line integral*

$$\int_L (\mathbf{a}, d\mathbf{r}) = \int_L P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz$$

to be independent of the integration path L it is necessary and sufficient for the vector field

$$\mathbf{a}(M) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

to be irrotational, i.e., for its curl to be zero throughout the field:

$$\operatorname{curl} \mathbf{a}(M) \equiv 0. \tag{24.95}$$

We assume here that the coordinates $P(x, y, z)$, $Q(x, y, z)$, and $R(x, y, z)$ of $\mathbf{a}(M)$ have continuous partial derivatives of the first order and the range of $\mathbf{a}(M)$ is a domain simply connected on the surface.

Remark. By virtue of Theorem 24.9, for a line integral to be independent of the integration path is equivalent to having a zero circulation of $\mathbf{a}$ around any closed contour. Therefore, we will prove the necessity and sufficiency of the condition (24.95) for the circulation to be zero.

◄ *Necessity.* Let a line integral be independent of the shape of the path, or, equivalently, let the circulation of $\mathbf{a}$ around any closed contour L be zero. Then

$$\operatorname{pr}_{\mathbf{n}} \operatorname{curl} \mathbf{a}\,|_M = \lim_{(\Sigma)\to M} \frac{\oint\limits_L (\mathbf{a}, d\mathbf{r})}{S} = 0.$$

i.e., at each point M of the field the projection of the curl $\mathbf{a}$ on any direction is zero.

This means that curl $\mathbf{a}$ is zero throughout the field, i.e.,

$$\operatorname{curl} \mathbf{a} \equiv 0.$$

Sufficiency. The sufficiency of (24.95) follows from Stokes' theorem, since if curl $\mathbf{a} \equiv 0$, then the circulation of $\mathbf{a}$ around any closed contour L is zero as well:

$$\oint\limits_L (\mathbf{a}, d\mathbf{r}) = \iint\limits_\Sigma (\operatorname{curl} \mathbf{a}, \mathbf{n}^0)\, d\sigma = 0. ►$$

For the plane field $\mathbf{a} = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ we have

$$\operatorname{curl} \mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P(x, y) & Q(x, y) & 0 \end{vmatrix} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}$$

and so for the plane field we have the following theorem:

Theorem 24.11. *For the line integral*

$$\int\limits_L P(x, y)\, dx + Q(x, y)\, dy \tag{24.96}$$

in a simply connected plane field to be independent of the shape of the line L, it is necessary and sufficient that the relation

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0 \tag{24.97}$$

hold identically throughout the entire domain in question.

If the domain is not simply connected, then generally speaking the condition

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \equiv 0$$

does not provide the independence of the line integral of the shape of the line.

Example. Let

$$\mathbf{a} = -\frac{y}{x^2 + y^2}\,\mathbf{i} + \frac{x}{x^2 + y^2}\,\mathbf{j}.$$

Consider the integral

$$\oint_L (\mathbf{a}, d\mathbf{r}) = \oint_L -\frac{y\,dx}{x^2 + y^2} + \frac{x\,dy}{x^2 + y^2}. \tag{24.98}$$

Clearly, the integrand has no sense at $O(0, 0)$. Therefore, we exclude this point. In the rest of the plane (which is not a simply connected domain!) the coordinates of $\mathbf{a}$ are continuous and have continuous partial derivatives and

$$\frac{\partial}{\partial y}\left(-\frac{y}{x^2 + y^2}\right) = \frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right).$$

Consider the integral (24.98) along the closed curve L, a circle of radius R with centre at the origin of coordinates

$$L: \begin{cases} x = R\cos t \\ y = R\sin t \end{cases} \qquad 0 \leqslant t \leqslant 2\pi.$$

Then

$$dx = -R\sin t\,dt, \qquad dy = R\cos t\,dt,$$

and

$$\oint_L \frac{x\,dy - y\,dx}{x^2 + y^2} = \int_0^{2\pi} \frac{R^2\cos^2 t + R^2\sin^2 t}{R^2}\,dt = \int_0^{2\pi} dt = 2\pi.$$

The fact that the circulation is nonzero indicates that the integral (24.98) depends on the shape of the integration path.

24.10 Potential Field

Definition. The field of a vector $\mathbf{a}(M)$ is called a *potential field*, if there exists a scalar function $u(M)$ such that

$$\operatorname{grad} u = \mathbf{a}. \tag{24.99}$$

Here $u(M)$ is called the *potential* of the field; its level surfaces are known as *equipotential surfaces*. Let

$$\mathbf{a} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}.$$

Since

$$\operatorname{grad} u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k},$$

the relation (24.99) is equivalent to three scalar equalities

$$\frac{\partial u}{\partial x} = P(x, y, z), \qquad \frac{\partial u}{\partial y} = Q(x, y, z), \qquad \frac{\partial u}{\partial z} = R(x, y, z).$$

Notice that the potential of the field is only determined up to a constant: if $\operatorname{grad} u = \mathbf{a}$ and $\operatorname{grad} v = \mathbf{a}$, then

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial x}, \qquad \frac{\partial u}{\partial y} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial z} = \frac{\partial v}{\partial z},$$

and hence $u = v + c$, where c is a constant.

Examples. (1) The field of the radius vector $\mathbf{r}$ is potential since

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = \operatorname{grad}\left(\frac{x^2 + y^2 + z^2}{2}\right) = \operatorname{grad}\frac{r^2}{2}.$$

Accordingly, the potential of the field of the radius vector is $r^2/2 + c$ (note that $r = \sqrt{x^2 + y^2 + z^2}$).

(2) The field of $\mathbf{a} = f(r)\mathbf{r}$ is also a potential field. We will show this. We find a function $\varphi(r)$ such that

$$f(r)\mathbf{r} = \operatorname{grad}\varphi(r).$$

We have $\operatorname{grad}\varphi(r) = \varphi'(r)\mathbf{r}^0$. Then

$$\varphi'(r)\mathbf{r}^0 = f(r)r\,\mathbf{r}^0,$$

whence $\varphi'(r) = f(r)r$. Therefore, $\varphi(r) = \int f(r)r\,dr$ is the potential of the field.

Theorem 24.12. *For the field of vector* $\mathbf{a}$ *to be a potential field it is necessary and sufficient for it to be irrotational, i.e., for its curl to be zero throughout the field*:

$$\operatorname{curl}\mathbf{a} \equiv 0. \tag{24.100}$$

This supposes that all the partial derivatives of the coordinates of **a** are continuous and the domain where **a** is defined is simply connected on the surface.

◄ Necessity. The necessity of (24.100) is established by direct calculations: if a field is potential, i.e., $\mathbf{a} = \mathbf{grad}\, u$, then

$$\mathbf{curl\, a} = \mathbf{curl}\,(\mathbf{grad}\, u) = \mathbf{curl}\left(\frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}\right)$$

$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial z} \end{vmatrix} = \left(\frac{\partial^2 u}{\partial z\,\partial y} - \frac{\partial^2 u}{\partial y\,\partial z}\right)\mathbf{i}$$

$$+ \left(\frac{\partial^2 u}{\partial x\,\partial y} - \frac{\partial^2 u}{\partial z\,\partial x}\right)\mathbf{j} + \left(\frac{\partial^2 u}{\partial y\,\partial x} - \frac{\partial^2 u}{\partial x\,\partial y}\right)\mathbf{k} = 0,$$

because mixed derivatives are independent of the order of differentiation.

Sufficiency. Let the field of vector **a** be irrotational, i.e., $\mathbf{curl\, a} \equiv 0$. Prove that the field is potential by actually constructing the potential $u(M)$ of the field.

Above all, from (24.100) we find that the line integral

$$\int_L (\mathbf{a},\, d\mathbf{r}) \tag{24.101}$$

is independent of the shape of the line L, and is only dependent on the positions of the initial and terminal points of the integration path. We fix the initial point $M_0(x_0, y_0, z_0)$ and will change the position of the terminal point $M(x, y, z)$. The integral (24.101) will then be a function of $M(x, y, z)$.

We denote this function by $u(M)$ and prove that it is the potential of the field, i.e., that

$$\mathbf{grad}\, u = \mathbf{a}.$$

We will write the integral (24.101) without indicating L but we will provide the initial and the finite points instead:

$$u(M) = \int_{M_0}^{M} (\mathbf{a},\, d\mathbf{r}) = \int_{(x_0,y_0,z_0)}^{(x,y,z)} P\,dx + Q\,dy + R\,dz. \tag{24.102}$$

The equality $\mathbf{grad}\, u = \mathbf{a}$ is equivalent to the three scalar relations

$$\frac{\partial u}{\partial x} = P(x, y, z), \qquad \frac{\partial u}{\partial y} = Q(x, y, z), \qquad \frac{\partial u}{\partial z} = R(x, y, z).$$

Prove the first of these, i.e., that

$$\frac{\partial u}{\partial x} = P(x, y, z).$$

The second and third ones are proved similarly.

By the definition of the partial derivative

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \to 0} \frac{u(x + \Delta x, y, z) - u(x, y, z)}{\Delta x}. \tag{24.103}$$

Consider the point $M_1(x + \Delta x, y, z)$ that is located close to $M(x, y, z)$. Since $u(M)$ is given by (24.102), where the line integral is independent of the integration path, we will choose the integration path as shown in Fig. 24.38.

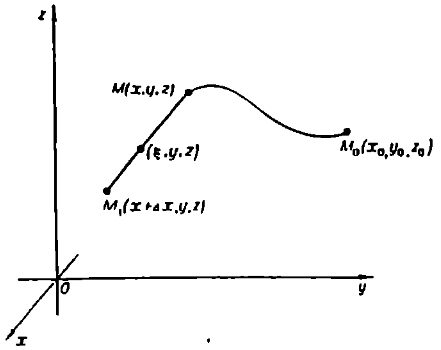


Fig. 24.38

Then

$$u(M_1) = \int_{M_0}^{M_1} (\mathbf{a}, d\mathbf{r}) = \int_{M_0}^{M} (\mathbf{a}, d\mathbf{r}) + \int_{M}^{M_1} (\mathbf{a}, d\mathbf{r}) = u(M) + \int_{M}^{M_1} (\mathbf{a}, d\mathbf{r}).$$

Hence

$$\Delta_x u = u(x + \Delta x, y, z) - u(x, y, z) = u(M_1) - u(M) = \int_{M}^{M_1} (\mathbf{a}, d\mathbf{r}).$$

The last integral is taken along the segment MM_1, which is parallel to the x-axis. On this segment we can choose as a parameter the coordinate x: $x = x$, $y =$ cosnt, $z =$ const, then $dx = dx$, $dy = 0$, $dz = 0$, so that

$$\Delta_x u = u(M_1) - u(M) = \int_{M(x,y,z)}^{M_1(x+\Delta x,y,z)} (\mathbf{a}, d\mathbf{r}) = \int_{x}^{x+\Delta x} P(x, y, z)\, dx. \tag{24.104}$$

We apply to the integral on the right of (24.104) the mean value theorem to find

$$\Delta_x u = P(\xi, y, z)\,\Delta x, \tag{24.105}$$

where ξ is confined between x and $x + \Delta x$. Further, from (24.105) we find

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta_x u}{\Delta x} = \lim_{\Delta x \to 0} P(\xi, y, z).$$

As $\Delta x \to 0$, the quantity $\xi \to x$, and since $P(x, y, z)$ is continuous, we obtain

$$\frac{\partial u}{\partial x} = P(x, y, z).$$

Similarly, we prove that

$$\frac{\partial u}{\partial y} = Q(x, y, z), \qquad \frac{\partial u}{\partial z} = R(x, y, z).$$

Corollary. A vector field is a potential field if and only if a line integral in it is independent of the path, i.e., if the circulation of the vector field around any closed contour in the field is zero.

Computing the line integral in a potential field. We will need the following theorem.

Theorem 24.13. *In a potential field* $\mathbf{a}(M)$ *the integral* $\int\limits_{M_1}^{M_2} (\mathbf{a}, d\mathbf{r})$ *is equal to the difference of the values of the potential* $u(M)$ *of the field at the terminal and the initial points of the integration path*

$$\int\limits_{M_1}^{M_2} (\mathbf{a}, d\mathbf{r}) = u(M_2) - u(M_1).$$

◄ We have proved above that the function

$$u(M) = \int\limits_{M_0}^{M} (\mathbf{a}, d\mathbf{r}) \tag{24.106}$$

is the potential of the field.

In a potential field the line integral $\int\limits_{M_1}^{M_2} (\mathbf{a}, d\mathbf{r})$ is independent of the shape of the integration path. Therefore, when choosing the path from M_1 to M_2 so that it passes through M_0 (Fig. 24.39) we will get

$$\int\limits_{M_1}^{M_2} (\mathbf{a}, d\mathbf{r}) = \int\limits_{M_1}^{M_0} (\mathbf{a}, d\mathbf{r}) + \int\limits_{M_0}^{M_2} (\mathbf{a}, d\mathbf{r})$$

or, changing the orientation of the path in the first integral on the right,

$$\int_{M_1}^{M_2} (\mathbf{a},\ d\mathbf{r}) = \int_{M_0}^{M_2} (\mathbf{a},\ d\mathbf{r}) - \int_{M_0}^{M_1} (\mathbf{a},\ d\mathbf{r}) = u(M_2) - u(M_1). \qquad (24.107)$$

Since the potential is determined up to a constant term, then any potential of the field $v(M)$ can be written as

$$v(M) = u(M) + C, \qquad (24.108)$$

where C is a constant.

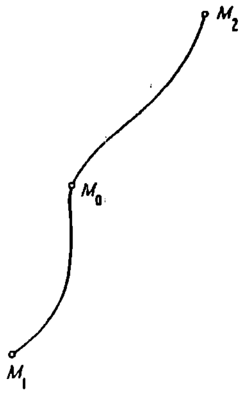


Fig. 24.39

Making replacements in (24.107): $u(M_2) = v(M_2) - C$, $u(M_1) = v(M_1) - C$, we will get for any $v(M)$

$$\int_{M_1}^{M_2} (\mathbf{a},\ d\mathbf{r}) = v(M_2) - v(M_1). ►$$

Example. We have shown that the potential of the field of the radius vector $\mathbf{r}$ is the function $u(r) = r^2/2$, therefore

$$\int_{M_1}^{M_2} (\mathbf{r},\ d\mathbf{r}) = \frac{1}{2}(r_2^2 - r_1^2),$$

where r_i $(i = 1, 2)$ is the distance of the points M_i $(i = 1, 2)$ from the origin of coordinates.

Computing the potential in rectangular coordinates. Let

$$\mathbf{a}(M) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}.$$

We have shown that the potential function $u(M)$ can be found from the formula

$$u(x, y, z) = \int_{M_0}^{M} P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz. \tag{24.109}$$

Integral (24.109) can more conveniently be found as follows. We fix the initial point $M_0(x_0, y_0, z_0)$ and connect it with a sufficiently close running point $M(x, y, z)$ by the broken line $M_0M_1M_2M$ whose legs are parallel to the coordinate axes, namely $M_0M_1 \parallel Ox$, $M_1M_2 \parallel Oy$, $M_2M \parallel Oz$ (Fig. 24.40).

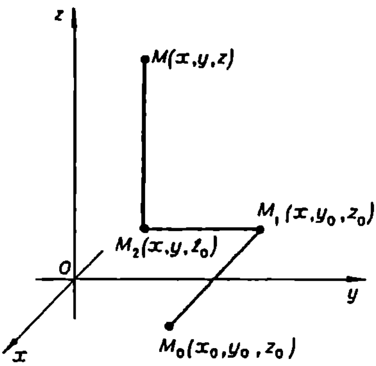


Fig. 24.40

On each leg only one coordinate varies, which enables us to significantly simplify our calculations. So, we have in the segment M_0M_1:

$$\begin{cases} x = x_0, & dx = dx, \\ y = y_0, & dy = 0, \\ z = z_0, & dz = 0. \end{cases}$$

In the segment M_1M_2:

$$\begin{cases} x = \text{const}, & dx = 0, \\ y = y, & dy = dy, \\ z = z_0, & dz = 0. \end{cases}$$

In the segment M_2M:

$$\begin{cases} x = \text{const}, & dx = 0, \\ y = \text{const}, & dy = 0, \\ z = z, & dz = dz. \end{cases}$$

And so the potential $u(M)$ will be

$$u(M) = \int_{M_0}^{M} (\mathbf{a}, d\mathbf{r}) = \int_{M_0}^{M_1} (\mathbf{a}, d\mathbf{r}) + \int_{M_1}^{M_2} (\mathbf{a}, d\mathbf{r}) + \int_{M_2}^{M} (\mathbf{a}, d\mathbf{r})$$

$$= \int_{x_0}^{x} P(x, y_0, z_0)\, dx + \int_{y_0}^{y} Q(x, y, z_0)\, dy + \int_{z_0}^{z} R(x, y, z)\, dz, \tag{24.110}$$

where x, y, z are the coordinates of the running point on the legs of the broken line, along which we integrate.

Example. Prove that the vector field

$$\mathbf{a} = (y + z)\mathbf{i} + (x + z)\mathbf{j} + (x + y)\mathbf{k}$$

is a potential field and find its potential.

◄ We check whether or not the field of $\mathbf{a}(M)$ is potential. To this end, we compute its curl:

$$\text{curl}\, \mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ y + z & x + z & x + y \end{vmatrix} = 0.$$

The field is thus a potential one.

We will find its potential using (24.110). We will take the origin of coordinates to be our initial point M_0, i.e., $x_0 = 0$, $y_0 = 0$, $z_0 = 0$. This is a normal procedure when a field $\mathbf{a}(M)$ is defined at the origin. We then obtain

$$u(x, y, z) = \int_0^x 0\, dx + \int_0^y (x + 0)\, dy + \int_0^z (x + y)\, dz = xy + (x + y)z.$$

Thus

$$u(x, y, z) = xy + xz + yz + C,$$

where C is an arbitrary constant.

There is another method. By definition, the potential $u(x, y, z)$ is a scalar function for which $\mathbf{grad}\, u = \mathbf{a}$. This vector equality is equivalent to the three scalar ones:

$$\frac{\partial u}{\partial x} = y + z, \tag{24.111}$$

$$\frac{\partial u}{\partial y} = x + z, \tag{24.112}$$

$$\frac{\partial u}{\partial z} = x + y. \tag{24.113}$$

Integrating (24.111) with respect to x gives

$$u(x, y, z) = \int_0^x (y + z)\,dx = xy + xz + f(y, z), \tag{24.114}$$

where $f(y, z)$ is an arbitrary differentiable function of y and z. Differentiating (24.114) with respect to y gives

$$\frac{\partial u}{\partial y} = x + \frac{\partial f(y, z)}{\partial y}.$$

From this we will have, by (24.112),

$$x + z = x + \frac{\partial f(y, z)}{\partial y},$$

or

$$\frac{\partial f(y, z)}{\partial y} = z. \tag{24.115}$$

Integrating (24.115) with respect to y gives

$$f(y, z) = \int_0^y z\,dy = yz + F(z), \tag{24.116}$$

where $F(z)$ is the yet undetermined function of z.

Substituing (24.116) into (24.114), we will obtain

$$u(x, y, z) = xy + xz + yz + F(z).$$

Differentiating this with respect to z we will obtain, by (24.113), an equation for $F(z)$

$$x + y = x + y + \frac{dF(z)}{dz}$$

whence $dF/dz = 0$, so that $F(z) = C = \text{const}$. Thus,

$$u(x, y, z) = xy + yz + zx + C. ►$$

24.11 Hamiltonian

We have discussed so far three basic operations of vector analysis: the construction of **grad** u for a scalar field $u = u(x, y, z)$ and the construction of **div a** and **curl a** for a vector field $\mathbf{a} = \mathbf{a}(x, y, z)$.

These operations can all be written in a more succinct form using the operator symbol

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}. \tag{24.117}$$

The operator ∇ (the *Hamiltonian operator* or *Hamiltonian*) has both differential and vector properties. We will agree that formal multiplication, e.g., of $\partial/\partial x$ by a function $u(x, y)$, means partial differentiation, i.e.,

$$\frac{\partial}{\partial x} \cdot u = \frac{\partial u}{\partial x}.$$

In vector algebra formal operations with ∇ are performed as if the operator is a vector. Using this formalism, we will obtain the following main relations:

(1) If $u = u(x, y, z)$ is a scalar differentiable function, then by the rule of multiplication of a vector by a scalar we will get

$$\nabla u = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k} = \mathbf{grad}\, u,$$

i.e.,

$$\nabla u = \mathbf{grad}\, u. \tag{24.118}$$

(2) If $\mathbf{a} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$, where P, Q, R are differentiable functions, then by the scalar product formula we will arrive at

$$\begin{aligned}(\nabla, \mathbf{a}) &= \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z},\ P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}\right)\\ &= \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = \operatorname{div}\mathbf{a},\end{aligned}$$

i.e.,

$$(\nabla, \mathbf{a}) = \operatorname{div}\mathbf{a}. \tag{24.119}$$

(3) The vector product $[\nabla, \mathbf{a}]$ gives

$$[\nabla, \mathbf{a}] = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} = \mathbf{curl}\, \mathbf{a}.$$

i.e.,

$$[\nabla, \mathbf{a}] = \mathbf{curl}\, \mathbf{a}. \tag{24.120}$$

For a constant function $u = C$ we clearly have $\nabla C = 0$, and for a constant vector $\mathbf{c}$ we will have

$$(\nabla, \mathbf{c}) = 0 \quad \text{and} \quad [\nabla, \mathbf{c}] = 0.$$

By the distributive property the scalar and vector products will be

$$(\nabla, \mathbf{a} + \mathbf{b}) = (\nabla, \mathbf{a}) + (\nabla, \mathbf{b}),$$

i.e.,

$$\operatorname{div}(\mathbf{a} + \mathbf{b}) = \operatorname{div}\mathbf{a} + \operatorname{div}\mathbf{b}, \qquad (24.121)$$

$$[\nabla, \mathbf{a} + \mathbf{b}] = [\nabla, \mathbf{a}] + [\nabla, \mathbf{b}],$$

i.e.,

$$\operatorname{curl}(\mathbf{a} + \mathbf{b}) = \operatorname{curl}\mathbf{a} + \operatorname{curl}\mathbf{b}. \qquad (24.122)$$

Remark. Formulas (24.121) and (24.122) can also be treated as a manifestation of the differential properties of the operator ∇, which is a linear differential operator. It was agreed that operator ∇ operates on all the quantities that follow it. So

$$(\nabla, \mathbf{a}) \neq (\mathbf{a}, \nabla),$$

because $(\nabla, \mathbf{a}) = \operatorname{div}\mathbf{a}$ is the function $\partial P/\partial x + \partial Q/\partial y + \partial R/\partial z$, while

$$(\mathbf{a}, \nabla) = P\frac{\partial}{\partial x} + Q\frac{\partial}{\partial y} + R\frac{\partial}{\partial z}$$

is a scalar differential operator.

Operating with ∇ on the product of some quantities one should remember the rules of differentiation of a product. It follows that one must apply operator ∇ to each factor in succession, while leaving the other factors unchanged. Then, one has to take the sum of the resultant expressions. In summary, one first takes into account the differential nature of the operator ∇, and then its vector properties.

Examples. (1) Prove that

$$\operatorname{\mathbf{grad}}(uv) = v \operatorname{\mathbf{grad}} u + u \operatorname{\mathbf{grad}} v. \qquad (24.123)$$

◄ Taking into account the above remark, we obtain from (24.118)

$$\nabla(uv) = v\nabla u + u\nabla v \quad \text{or} \quad \operatorname{\mathbf{grad}}(uv) = v \operatorname{\mathbf{grad}} u + u \operatorname{\mathbf{grad}} v. \text{ ►}$$

To indicate that ∇ does not operate on some quantity that enters into a complex formula one uses the subscript c (const), which in the final result is discarded.

(2) Let $u(x, y, z)$ be a scalar differentiable function and $\mathbf{a}(x, y, z)$ be a vector differentiable function.

Prove that

$$\operatorname{div}(u\mathbf{a}) = u \operatorname{div}\mathbf{a} + (\mathbf{a}, \operatorname{\mathbf{grad}} u). \qquad (24.124)$$

◄ We rewrite the left-hand side of (24.124) as

$$\operatorname{div}(u\mathbf{a}) = (\nabla, u\mathbf{a}).$$

At first we take into account the differential nature of the operator ∇; we get

$$(\nabla, u\mathbf{a}) = (\nabla, u_c\mathbf{a}) + (\nabla, u\mathbf{a}_c).$$

Since u_c is a constant scalar, it can be placed outside the sign of the scalar product, so that

$$(\nabla, u_c\mathbf{a}) = u_c(\nabla, \mathbf{a}) = u_c \operatorname{div} \mathbf{a} = u \operatorname{div} \mathbf{a}$$

(after the last operation we discarded the subscript c).

In $(\nabla, u\mathbf{a}_c)$ the operator ∇ operates only on the scalar function u, therefore

$$(\nabla, u\mathbf{a}_c) = (\nabla u, \mathbf{a}_c) = (\mathbf{a}_c, \nabla u) = (\mathbf{a}, \operatorname{grad} u).$$

Thus

$$\operatorname{div}(u\mathbf{a}) = u \operatorname{div} \mathbf{a} + (\mathbf{a}, \operatorname{grad} u). \blacktriangleright$$

Remark. Using the formalism of operations with ∇ as a vector, one should remember that ∇ is no vector, it has neither magnitude nor direction, so that, for example, the vector $[\nabla, \mathbf{a}]$ will, generally speaking, be not perpendicular to $\mathbf{a}$ (however, for the plane field $\mathbf{a} = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ the vector

$$[\nabla, \mathbf{a}] = \operatorname{curl} \mathbf{a} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}$$

will be perpendicular to the xy-plane, and hence to the vector $\mathbf{a}$ as well).

Further, the notion of collinearity makes no sense for the symbolic vector ∇. For example, the expression $[\nabla\varphi, \nabla\psi]$, where φ and ψ are scalar functions, formally resembles the vector product of two collinear vectors, which is always zero. In the general case, however, this is not the case. In fact, the vector $\nabla\varphi = \operatorname{grad} \varphi$ points along the normal to the level surface

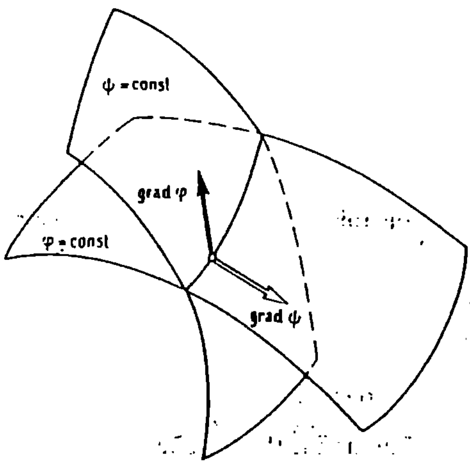


Fig. 24.41

φ = const and the vector $\nabla\psi$ = **grad** ψ determines the normal to the level surface ψ = const, and these normals in the general case should not necessarily be collinear (Fig. 24.41). On the other hand, in any differentiable scalar field $\varphi(x, y, z)$ we have $[\nabla\varphi, \nabla\psi] = 0$.

The examples just considered show that the operator ∇ should be treated with great care, and should one be dubuous about the end result one might be recommended to test it by analytic means.

24.12 Differential Operations of the Second Order. Laplace Operator

Differential operations of the second order are a result of double application of the operator ∇ to fields.

(1) Consider the scalar field $u = u(x, y, z)$. In this filed the operator ∇ gives rise to a vector field

$$\nabla u = \mathbf{grad}\, u.$$

In the vector field **grad** u we can define two operations:

$$(\nabla, \nabla u) = \operatorname{div} \mathbf{grad}\, u, \tag{24.125}$$

which leads to a scalar field, and

$$[\nabla, \nabla u] = \mathbf{curl}\, \mathbf{grad}\, u, \tag{24.126}$$

which leads to a vector field.

(2) Suppose we have a vector field $\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$. The operator ∇ gives rise in it to the scalar field $(\nabla, \mathbf{a}) = \operatorname{div} \mathbf{a}$. In the scalar field $\operatorname{div} \mathbf{a}$ the operator ∇ produces a vector field

$$\nabla(\nabla, \mathbf{a}) = \mathbf{grad} \operatorname{div} \mathbf{a}. \tag{24.127}$$

(3) In the vector field $\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ the operator ∇ produces a vector field as well:

$$[\nabla, \mathbf{a}] = \mathbf{curl}\, \mathbf{a}.$$

If again we operate on this field with ∇, we will get:

(a) a scalar field

$$(\nabla, [\nabla, \mathbf{a}]) = \operatorname{div} \mathbf{curl}\, \mathbf{a}, \tag{24.128}$$

(b) a vector field

$$[\nabla, [\nabla, \mathbf{a}]] = \mathbf{curl}\, \mathbf{curl}\, \mathbf{a}. \tag{24.129}$$

Formulas (24.125-129) define the so-called *differential operations of the second order.*

Now we select in space a rectangular coordinate system $Oxyz$ and consider in more detail each of the formulas (24.125-129).

(1) Assuming that the function $u(x, y, z)$ has second partial derivatives with respect to x, y and z, we will obtain

$$\begin{aligned}\operatorname{div}\mathbf{grad}\, u &= (\nabla, \nabla u) = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}, \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}\right) \\ &= \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial u}{\partial z}\right) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.\end{aligned}$$

The symbol

$$\Delta \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

is known as the *Laplace operator*, or *Laplacian*. It can be represented as the scalar product of the Hamiltonian ∇ by itself, i.e.,

$$\Delta = (\nabla, \nabla) = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{24.130}$$

The operator Δ plays an important role in mathematical physics. The equation $\Delta u = 0$ or

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

is called the *Laplace equation*. It describes, for instance, the stationary propagation of heat.

A scalar field $u(x, y, z)$ meeting the condition $\Delta u = 0$ is said to be the *Laplace* or *harmonic field*. For example, the scalar field $u = 2x^2 + 3y - 2z^2$ is harmonic throughout the entire three-dimensional space. We will see that this is so.

◄ Since

$$\frac{\partial u}{\partial x} = 4x, \quad \frac{\partial u}{\partial y} = 3, \quad \frac{\partial u}{\partial z} = -4z,$$
$$\frac{\partial^2 u}{\partial x^2} = 4, \quad \frac{\partial^2 u}{\partial y^2} = 0, \quad \frac{\partial^3 u}{\partial z^2} = -4,$$

then

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 4 + 0 - 4 = 0$$

which proves that the field is harmonic. ►

(2) Let $u(x, y, z)$ have continuous partial derivatives of the second order. Then

$$\mathbf{curl}\,\mathbf{grad}\, u \equiv 0. \tag{24.131}$$

Formally, we will obtain

$$\operatorname{curl}\operatorname{grad} u = [\nabla, \nabla u] = [\nabla, \nabla]u = 0,$$

because $[\nabla, \nabla] = 0$ as a vector product of two identical "vectors".

The same result can be obtained using the expressions for the gradient and curl in rectangular coordinates and taking into account the conditions of the formulation:

$$\operatorname{curl}\operatorname{grad} u = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \end{vmatrix} = \left(\frac{\partial^2 u}{\partial z\,\partial y} - \frac{\partial^2 u}{\partial y\,\partial z}\right)\mathbf{i}$$

$$+ \left(\frac{\partial^2 u}{\partial z\,\partial x} - \frac{\partial^2 u}{\partial x\,\partial z}\right)\mathbf{j} + \left(\frac{\partial^2 u}{\partial y\,\partial x} - \frac{\partial^2 u}{\partial x\,\partial y}\right)\mathbf{k} = 0.$$

We have

$$\frac{\partial^2 u}{\partial z\,\partial y} = \frac{\partial^2 u}{\partial y\,\partial z}, \qquad \frac{\partial^2 u}{\partial x\,\partial z} = \frac{\partial^2 u}{\partial z\,\partial x}, \qquad \frac{\partial^2 u}{\partial y\,\partial x} = \frac{\partial^2 u}{\partial x\,\partial y}.$$

since the partial derivatives of the second order are continuous.

(3) Let

$$\mathbf{a} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

be a vector field, whose coordinates P, Q, R have continuous partial derivatives of the second order. We then obtain

$$\begin{aligned} \operatorname{grad}\operatorname{div}\mathbf{a} = \nabla(\nabla, \mathbf{a}) &= \frac{\partial}{\partial x}\left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)\mathbf{i} \\ &+ \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)\mathbf{j} + \frac{\partial}{\partial z}\left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)\mathbf{k} \\ &= \left(\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 Q}{\partial y\,\partial x} + \frac{\partial^2 R}{\partial z\,\partial x}\right)\mathbf{i} + \left(\frac{\partial^2 P}{\partial x\,\partial y} + \frac{\partial^2 Q}{\partial y^2} + \frac{\partial^2 R}{\partial z\,\partial y}\right)\mathbf{j} \\ &+ \left(\frac{\partial^2 P}{\partial x\,\partial z} + \frac{\partial^2 Q}{\partial y\,\partial z} + \frac{\partial^2 R}{\partial z^2}\right)\mathbf{k}. \end{aligned} \tag{24.132}$$

(4) Under the same conditions as in (3) we will have

$$\operatorname{div}\operatorname{curl}\mathbf{a} = 0. \tag{24.133}$$

This relation has already been proved earlier directly (see (24.82′)). We will provide here a formal proof of relation (24.133) using the well-known formula of vector algebra

$$(\mathbf{A}, [\mathbf{B}, \mathbf{C}]) = (\mathbf{C}, [\mathbf{A}, \mathbf{B}]) = (\mathbf{B}, [\mathbf{C}, \mathbf{A}]).$$

We have

$$\operatorname{div}\operatorname{curl}\mathbf{a} = (\nabla, [\nabla, \mathbf{a}]) = (\mathbf{a}, [\nabla, \nabla]) = 0,$$

since $[\nabla, \nabla] = 0$ as a vector product of two identical "vectors".

(5) Show, lastly, that under the same conditions as before, we have

$$\operatorname{curl}\operatorname{curl}\mathbf{a} = \operatorname{grad}\operatorname{div}\mathbf{a} - \Delta a. \tag{24.134}$$

In fact, since

$$\operatorname{curl}\operatorname{curl}\mathbf{a} = [\nabla, [\nabla, \mathbf{a}]],$$

then by the formula of double vector product

$$[\mathbf{A}, [\mathbf{B}, \mathbf{C}]] = \mathbf{B}(\mathbf{A}, \mathbf{C}) - (\mathbf{A}, \mathbf{B})\mathbf{C},$$

putting in it $\mathbf{A} = \nabla$, $\mathbf{B} = \nabla$, $\mathbf{C} = \mathbf{a}$, we will get

$$[\nabla, [\nabla, \mathbf{a}]] = \nabla(\nabla, \mathbf{a}) - (\nabla, \nabla)\mathbf{a}.$$

But $(\nabla, \mathbf{a}) = \operatorname{div}\mathbf{a}$, and $(\nabla, \nabla) = \Delta$. Therefore, we will arrive at

$$\operatorname{curl}\operatorname{curl}\mathbf{a} = \operatorname{grad}\operatorname{div}\mathbf{a} - \Delta\mathbf{a},$$

where grad div a is given by (24.132), and $\Delta\mathbf{a}$ for $\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ should be understood to be

$$\Delta\mathbf{a} = \Delta P\,\mathbf{i} + \Delta Q\,\mathbf{j} + \Delta R\,\mathbf{k}.$$

Before we leave the section we will provide a table of differential operations of the second order

	Scalar field $u = u(x, y, z)$	Vector field $\mathbf{a} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$	
	grad	**div**	**curl**
grad		**grad div a**	
div	**div grad** $u = \Delta u$		**div curl a** $= 0$
curl	**curl grad** $u = 0$		**curl curl a** $=$ **grad div a** $- \Delta\mathbf{a}$

The empty rectangles imply that the operation in question makes no sense (e.g., **grad curl a**).

24.13 Curvilinear Coordinates

In many applications it is more convenient to define the position of a point in space not by three rectangular coordinates (x, y, z) but by some other numbers (q_1, q_2, q_3) which are more suitable in a given case.

Suppose that corresponding to each point M in space is a set of numbers (q_1, q_2, q_3); conversely, corresponding to each such set is a single point M. The quantities q_1, q_2, q_3 are then called the *curvilinear coordinates* of the point M.

The coordinate surfaces in a system of curvilinear coordinates q_1, q_2, q_3 are the surfaces $q_1 = C_1$, $q_2 = C_2$, $q_3 = C_3$.

On the *coordinate surfaces* one of the coordinates remains constant.

The lines of intersection of two coordinate surfaces are called the *coordinate lines*.

Examples of curvilinear coordinates are cylindrical and spherical coordinates.

Cylindrical coordinates. In cylindrical coordinates the position of a point M in space is defined by the three coordinates

$$\begin{cases} q_1 = \varrho, & 0 \leqslant \varrho < +\infty, \\ q_2 = \varphi, & 0 \leqslant \varphi < 2\pi, \\ q_3 = z, & -\infty < z < +\infty. \end{cases} \tag{24.135}$$

The coordinate surfaces are (Fig. 24.42):

ϱ = const — circular cylinders with the z-axis as their axis;

φ = const — half-planes adjacent to the z-axis;

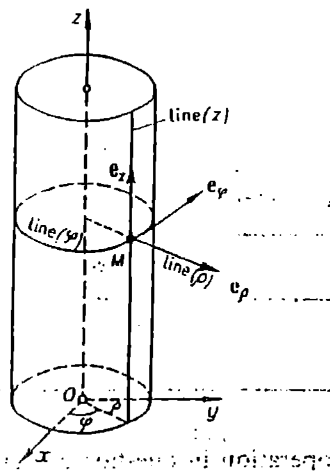


Fig. 24.42

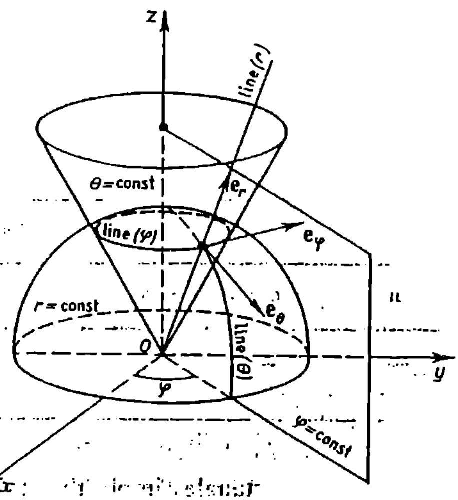


Fig. 24.43

z = const — planes perpendicular to the z-axis.

The coordinate lines are:

ϱ-lines — rays perpendicular to the z-axis originating on the z-axis;

φ-lines — circles with centre on the z-axis lying in planes perpendicular to the z-axis;

z-lines — straight lines parallel to the z-axis.

The rectangular coordinates (x, y, z) of a point are related to its cylindrical coordinates (ϱ, φ, z) by

$$\begin{cases} x = \varrho \cos \varphi, \\ y = \varrho \sin \varphi, \\ z = z. \end{cases} \tag{24.136}$$

Spherical coordinates. In spherical coordinates the position of a point M in space is defined by

$$\begin{cases} q_1 = r, & 0 \leqslant r < +\infty, \\ q_2 = \theta, & 0 \leqslant \theta \leqslant \pi, \\ q_3 = \varphi, & 0 \leqslant \varphi < 2\pi. \end{cases} \tag{24.137}$$

The coordinate surfaces are (Fig. 24.43):

r = const — spheres with centre at point O;

θ = const — circular half-cones with the z-axis as their axes;

φ = const — half-planes adjacent to the z-axis.

The coordinate lines are:

r-lines — rays originating at point O;

θ-lines — meridians on a sphere;

φ-lines — parallels on a sphere.

The rectangular coordinates (x, y, z) of a point are related to its spherical coordinates (r, θ, φ) by

$$\begin{cases} x = r \cos \varphi \sin \theta, \\ y = r \sin \varphi \sin \theta, \\ z = r \cos \theta. \end{cases} \tag{24.138}$$

We introduce the unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, directed along tangents to the coordinate lines (q_1), (q_2) and (q_3) at the point M in the direction of the growth of q_1, q_2 and q_3, respectively.

Definition. A system of curvilinear coordinates is called *orthogonal*, if at each point M the unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ are pairwise orthogonal.

In such a system the coordinate lines and the coordinate surfaces will also be orthogonal.

Examples of orthogonal curvilinear coordinate systems are systems of cylindrical and spherical coordinates. We will confine ourselves to orthogonal systems of coordinates.

Let $\mathbf{r} = \mathbf{r}(q_1, q_2, q_3)$ be the radius vector of the point M. We can then show that

$$d\mathbf{r} = H_1\, dq_1\, \mathbf{e}_1 + H_2\, dq_2\, \mathbf{e}_2 + H_3\, dq_3\, \mathbf{e}_3 \tag{24.139}$$

where

$$H_i = \sqrt{\left(\frac{\partial x}{\partial q_i}\right)^2 + \left(\frac{\partial y}{\partial q_i}\right)^2 + \left(\frac{\partial z}{\partial q_i}\right)^2} \qquad (i = 1,\ 2,\ 3)$$

are the Lamé coefficients for a given system of curvilinear coordinates. We have:

$$q_1 = \varrho, \qquad q_2 = \varphi, \qquad q_3 = z$$

and

$$x = \varrho \cos\varphi, \qquad y = \varrho \sin\varphi, \qquad z = z.$$

Hence

$$\begin{aligned} H_1 &= H_\varrho = \sqrt{\left(\frac{\partial x}{\partial \varrho}\right)^2 + \left(\frac{\partial y}{\partial \varrho}\right)^2 + \left(\frac{\partial z}{\partial \varrho}\right)^2} = 1, \\ H_2 &= H_\varphi = \sqrt{\left(\frac{\partial x}{\partial \varphi}\right)^2 + \left(\frac{\partial y}{\partial \varphi}\right)^2 + \left(\frac{\partial z}{\partial \varphi}\right)^2} = \varrho, \\ H_3 &= H_z = \sqrt{\left(\frac{\partial x}{\partial z}\right)^2 + \left(\frac{\partial y}{\partial z}\right)^2 + \left(\frac{\partial z}{\partial z}\right)^2} = 1. \end{aligned} \tag{24.140}$$

Likewise, for spherical coordinates we have

$$H_1 = H_r = 1, \qquad H_2 = H_\theta = r, \qquad H_3 = H_\varphi = r \sin\theta. \tag{24.141}$$

The quantities

$$dl_1 = H_1\, dq_1, \qquad dl_2 = H_2\, dq_2, \qquad dl_3 = H_3\, dq_3 \tag{24.142}$$

are the differentials of arcs of the corresponding coordinate lines.

24.14 Basic Vector Operations in Curvilinear Coordinates

Differential equations of vector lines. Suppose that we have the field of a vector

$$\mathbf{a} = a_1(q_1, q_2, q_3)\mathbf{e}_1 + a_2(q_1, q_2, q_3)\mathbf{e}_2 + a_3(q_1, q_2, q_3)\mathbf{e}_3.$$

In the curvilinear coordinates q_1, q_2 and q_3 the equations of the vector lines are

$$\frac{H_1\, dq_1}{a_1(q_1, q_2, q_3)} = \frac{H_2\, dq_2}{a_2(q_1, q_2, q_3)} = \frac{H_3\, dq_3}{a_3(q_1, q_2, q_3)}.$$

Specifically, in cylindrical coordinates ($q_1 = \varrho, q_2 = \varphi, q_3 = z$) we have

$$\frac{d\varrho}{a_1(\varrho, \varphi, z)} = \frac{\varrho\, d\varphi}{a_2(\varrho, \varphi, z)} = \frac{dz}{a_3(\varrho, \varphi, z)}, \tag{24.143}$$

and in spherical coordinates ($q_1 = r, q_2 = \theta, q_3 = \varphi$) we have

$$\frac{dr}{a_1(r, \theta, \varphi)} = \frac{r\, d\theta}{a_2(r, \theta, \varphi)} = \frac{r \sin\theta\, d\varphi}{a_3(r, \theta, \varphi)}. \tag{24.144}$$

Gradient in orthogonal coordinates. Consider the scalar field

$$u = u(q_1, q_2, q_3).$$

Then

$$\operatorname{grad} u = \frac{1}{H_1}\frac{\partial u}{\partial q_1}\mathbf{e}_1 + \frac{1}{H_2}\frac{\partial u}{\partial q_2}\mathbf{e}_2 + \frac{1}{H_3}\frac{\partial u}{\partial q_3}\mathbf{e}_3.$$

In particular, in cylindrical coordinates ($q_1 = \varrho, q_2 = \varphi, q_3 = z$) we have

$$\operatorname{grad} u = \frac{\partial u}{\partial \varrho}\mathbf{e}_\varrho + \frac{1}{\varrho}\frac{\partial u}{\partial \varphi}\mathbf{e}_\varphi + \frac{\partial u}{\partial z}\mathbf{e}_z. \tag{24.145}$$

In spherical coordinates ($q_1 = r, q_2 = \theta, q_3 = \varphi$) we have

$$\operatorname{grad} u = \frac{\partial u}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial u}{\partial \theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial u}{\partial \varphi}\mathbf{e}_\varphi. \tag{24.146}$$

Curl in orthogonal coordinates. Consider the vector field

$$\mathbf{a} = a_1(q_1, q_2, q_3)\mathbf{e}_1 + a_2(q_1, q_2, q_3)\mathbf{e}_2 + a_3(q_1, q_2, q_3)\mathbf{e}_3.$$

Its curl will be

$$\operatorname{curl} \mathbf{a} = \begin{vmatrix} \dfrac{\mathbf{e}_1}{H_2H_3} & \dfrac{\mathbf{e}_2}{H_1H_3} & \dfrac{\mathbf{e}_3}{H_1H_2} \\ \dfrac{\partial}{\partial q_1} & \dfrac{\partial}{\partial q_2} & \dfrac{\partial}{\partial q_3} \\ a_1H_1 & a_2H_2 & a_3H_3 \end{vmatrix}.$$

Specifically, in cylindrical coordinates ($q_1 = \varrho, q_2 = \varphi, q_3 = z, H_1 = 1, H_2 = \varrho, H_3 = 1$) we have

$$\operatorname{curl} \mathbf{a} = \begin{vmatrix} \dfrac{1}{\varrho}\mathbf{e}_\varrho & \mathbf{e}_\varphi & \dfrac{1}{\varrho}\mathbf{e}_z \\ \dfrac{\partial}{\partial \varrho} & \dfrac{\partial}{\partial \varphi} & \dfrac{\partial}{\partial z} \\ a_1 & a_2\varrho & a_3 \end{vmatrix}. \tag{24.147}$$

In spherical coordinates ($q_1 = r$, $q_2 = \theta$, $q_3 = \varphi$, $H_1 = 1$, $H_2 = r$, $H_3 = r\sin\theta$) we have

$$\operatorname{curl} \mathbf{a} = \begin{vmatrix} \dfrac{\mathbf{e}_r}{r^2 \sin\theta} & \dfrac{\mathbf{e}_\theta}{r\sin\theta} & \mathbf{e}_\varphi \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \theta} & \dfrac{\partial}{\partial \varphi} \\ a_1 & a_2 r & a_3 r \sin\theta \end{vmatrix} . \tag{24.148}$$

Divergence in orthogonal coordinates. Consider the vector field

$$\mathbf{a} = a_1(q_1, q_2, q_3)\mathbf{e}_1 + a_2(q_1, q_2, q_2)\mathbf{e}_2 + a_3(q_1, q_2, q_3)\mathbf{e}_3.$$

Then $\operatorname{div} \mathbf{a}$ is computed by the formula

$$\operatorname{div} \mathbf{a} = \frac{1}{H_1 H_2 H_3}\left[\frac{\partial(a_1 H_2 H_3)}{\partial q_1} + \frac{\partial(a_2 H_1 H_3)}{\partial q_2} + \frac{\partial(a_3 H_1 H_2)}{\partial q_3}\right]. \tag{24.149}$$

In particular, in cylindrical coordinates ($q_1 = \varrho$, $q_2 = \varphi$, $q_3 = z$) we have

$$\operatorname{div} \mathbf{a} = \frac{1}{\varrho}\left[\frac{\partial(a_1\varrho)}{\partial\varrho} + \frac{\partial(a_2)}{\partial\varphi} + \frac{\partial(a_3\varrho)}{\partial z}\right],$$

or

$$\operatorname{div} \mathbf{a} = \frac{1}{\varrho}\cdot\frac{\partial(a_1\varrho)}{\partial\varrho} + \frac{1}{\varrho}\frac{\partial a_2}{\partial\varphi} + \frac{\partial a_3}{\partial z}.$$

In spherical coordinates ($q_1 = r$, $q_2 = \theta$, $q_3 = \varphi$) we have

$$\operatorname{div} \mathbf{a} = \frac{1}{r^2}\frac{\partial(a_1 r^2)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial(a_2\sin\theta)}{\partial\theta} + \frac{1}{r\sin\theta}\frac{\partial a_3}{\partial\varphi}.$$

Applying (24.149) to the unit vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ gives

$$\operatorname{div} \mathbf{e}_1 = \frac{1}{H_1 H_2 H_3}\frac{\partial(H_2 H_3)}{\partial q_1}$$

$$\operatorname{div} \mathbf{e}_2 = \frac{1}{H_1 H_2 H_3}\frac{\partial(H_3 H_1)}{\partial q_2}$$

$$\operatorname{div} \mathbf{e}_3 = \frac{1}{H_1 H_2 H_3}\frac{\partial(H_1 H_2)}{\partial q_3}$$

Flux in curvilinear coordinates. Let S be part of the coordinate surface $q_1 = C = \text{const}$ bounded by the coordinate lines

$$\begin{cases} q_2 = \alpha_1, & q_2 = \alpha_2 \quad (\alpha_1 < \alpha_2), \\ q_3 = \beta_1, & q_3 = \beta_2 \quad (\beta_1 < \beta_2). \end{cases}$$

Then the flux of a vector

$$\mathbf{a} = a_1(q_1, q_2, q_3)\mathbf{e}_1 + a_2(q_1, q_2, q_3)\mathbf{e}_2 + a_3(q_1, q_2, q_3)\mathbf{e}_3$$

through the surface S in the direction $\mathbf{e}_1$ will be given by

$$\Pi = \int_{\alpha_1}^{\alpha_2} \int_{\beta_1}^{\beta_2} a_1(C, q_2, q_3)H_2(C, q_2, q_3)H_3(C, q_2, q_3)dq_3dq_2. \tag{24.150}$$

In a similar manner we work out the flux through the parts $q_2 = C$ and $q_3 = C$, where $C = \text{const}$, of the surface.

Example. Find the flux Π of the vector field

$$\mathbf{a} = r^2\theta\mathbf{e}_r + re^{2\theta}\mathbf{e}_\theta$$

through the external side of the upper hemisphere S of radius R with centre at the origin of coordinates.

◄ The hemisphere S is part of the coordinate surface $r = \text{const}$, namely $r = R$. On the hemisphere S we have $q_1 = r = R$, i.e., $C = R$; $q_2 = \theta$, where $0 \leqslant \theta \leqslant \pi/2$ (i.e., $\alpha_1 = 0$, $\alpha_2 = \pi/2$); $q_3 = \varphi$, where $0 \leqslant \varphi \leqslant 2\pi$ (i.e., $\beta_1 = 0$, $\beta_2 = 2\pi$).

Considering that in spherical coordinates $H_1 = H_r = 1$, $H_2 = H_\theta = r$, $H_3 = H_\varphi = r\sin\theta$, we find by (24.150)

$$\Pi = \int_0^{\pi/2} d\theta \int_0^{2\pi} R^2\theta \sin\theta \, d\varphi = 2\pi R^4 \int_0^{\pi/2} \theta \sin\theta \, d\theta = 2\pi R^4. \text{ ►}$$

Potential in curvilinear coordinates. Let a potential vector field

$$\mathbf{a}(M) = a_1(q_1, q_2, q_3)\mathbf{e}_1 + a_2(q_1, q_2, q_3)\mathbf{e}_2 + a_3(q_1, q_2, q_3)\mathbf{e}_3$$

be defined in some domain Ω, i.e., $\operatorname{curl} \mathbf{a} = 0$ in Ω.

To find the potential $u(q_1, q_2, q_3)$ of $\mathbf{a}$, we write

$$\mathbf{a}(M) = \operatorname{grad} u(M)$$

in the form

$$a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 = \frac{1}{H_1}\frac{\partial u}{\partial q_1}\mathbf{e}_1 + \frac{1}{H_2}\frac{\partial u}{\partial q_2}\mathbf{e}_2 + \frac{1}{H_3}\frac{\partial u}{\partial q_3}\mathbf{e}_3.$$

It follows that

$$\frac{\partial u}{\partial q_1} = H_1a_1, \qquad \frac{\partial u}{\partial q_2} = H_2a_2, \qquad \frac{\partial u}{\partial q_3} = H_3a_3. \tag{24.151}$$

Integrating the system of differential equations with partial derivatives (24.151), we will find the potential

$$u = u(q_1, q_2, q_3) + C,$$

where C is an arbitrary constant.

Notice that the system (24.151) is solved in the same manner as in finding the potential in rectangular coordinates. Specifically, in cylindrical coordinates ($q_1 = \varrho$, $q_2 = \varphi$, $q_3 = z$, $H_1 = 1$, $H_2 = \varrho$, $H_3 = 1$) system (24.151) becomes

$$\frac{\partial u}{\partial \varrho} = a_1, \qquad \frac{\partial u}{\partial \varphi} = a_2 \varrho, \qquad \frac{\partial u}{\partial z} = a_3, \tag{24.152}$$

where a_1, a_2, a_3 are the cylindrical coordinates of the vector $\mathbf{a}(M)$, i.e.,

$$\mathbf{a} = a_1(\varrho, \varphi, z)\mathbf{e}_\varrho + a_2(\varrho, \varphi, z)\mathbf{e}_\varphi + a_3(\varrho, \varphi, z)\mathbf{e}_z.$$

In spherical coordinates ($q_1 = r$, $q_2 = \theta$, $q_3 = \varphi$, $H_1 = 1$, $H_2 = r$, $H_3 = r \sin\theta$) system (24.151) becomes

$$\frac{\partial u}{\partial r} = a_1, \qquad \frac{\partial u}{\partial \theta} = a_2 r, \qquad \frac{\partial u}{\partial \varphi} = a_3 r \sin\theta,$$

where

$$\mathbf{a}(M) = a_1(r, \theta, \varphi)\mathbf{e}_r + a_3(r, \theta, \varphi)\mathbf{e}_\theta + a_3(r, \theta, \varphi)\mathbf{e}_\varphi.$$

Example. Find the potential of a vector field specified in cylindrical coordinates

$$\mathbf{a} = \left(\frac{\tan^{-1} z}{\varrho} + \cos\varphi\right)\mathbf{e}_\varrho - \sin\varphi\, \mathbf{e}_\varphi + \frac{\ln \varrho}{1 + z^2}\mathbf{e}_z.$$

◄ We will see that $\operatorname{curl} \mathbf{a} = 0$. By (24.147),

$$\operatorname{curl} \mathbf{a} = \begin{vmatrix} \dfrac{\mathbf{e}_\varrho}{\varrho} & \mathbf{e}_\varphi & \dfrac{\mathbf{e}_z}{\varrho} \\ \dfrac{\partial}{\partial \varrho} & \dfrac{\partial}{\partial \varphi} & \dfrac{\partial}{\partial z} \\ \dfrac{\tan^{-1} z}{\varrho} + \cos\varphi & -\varrho \sin\varphi & \dfrac{\ln \varrho}{1 + z^2} \end{vmatrix} = 0,$$

i.e., the field is a potential one.

The potential $u = u(\varrho, \varphi, z)$ is a solution of the following system of differential equations with partial derivatives (see (24.152))

$$\begin{cases} \dfrac{\partial u}{\partial \varrho} = \dfrac{\tan^{-1} z}{\varrho} + \cos\varphi, \\ \dfrac{\partial u}{\partial \varphi} = -\varrho \sin\varphi, \\ \dfrac{\partial u}{\partial z} = \dfrac{\ln \varrho}{1 + z^2}. \end{cases}$$

Integrating with respect to ϱ, we find from the first of these equations

$$u = \tan^{-1} z \ln \varrho + \varrho \cos\varphi + C(\varphi, z). \tag{24.153}$$

Differentiating (24.153) with respect to φ and using the second equation, we obtain

$$-\varrho \sin\varphi + \frac{\partial C(\varphi, z)}{\partial \varphi} = -\varrho \sin\varphi$$

or

$$\frac{\partial C(\varphi, z)}{\partial \varphi} = 0,$$

hence $C = C_1(z)$. Thus

$$u = \tan^{-1} z \ln \varrho + \varrho \cos\varphi + C_1(z).$$

Differentiating this with respect to z and using the third equation, we get

$$\frac{\ln \varrho}{1 + z^2} + C_1'(z) = \frac{\ln \varrho}{1 + z^2},$$

or

$$C'(z) = 0,$$

hence $C_1(z) = C$. Accordingly, the potential of the field will be

$$u(\varrho, \varphi, z) = \tan^{-1} z \ln \varrho + \varrho \cos\varphi + C. \blacktriangleright$$

Line integral and circulation in curvilinear coordinates. Let a vector field

$$\mathbf{a}(M) = a_1(q_1, q_2, q_3)\mathbf{e}_1 + a_2(q_1, q_2, q_3)\mathbf{e}_2 + a_3(q_1, q_2, q_3)\mathbf{e}_3$$

be defined and continuous in a domain Ω of orthogonal curvilinear coordinates q_1, q_2, q_3. Since the differential of a radius vector $\mathbf{r}$ of any point $M(q_1, q_2, q_3) \in \Omega$ is given by

$$d\mathbf{r} = H_1\, dq_1\, \mathbf{e}_1 + H_2\, dq_2\, \mathbf{e}_2 + H_3\, dq_3\, \mathbf{e}_3, \tag{24.154}$$

then the line integral of $\mathbf{a}(M)$ along an oriented smooth or piecewise smooth curve $L \subset \Omega$ will be

$$\int_L (\mathbf{a}, d\mathbf{r}) = \int_L a_1H_1dq_1 + a_2H_2dq_2 + a_3H_3dq_3. \tag{24.155}$$

Specifically, for cylindrical coordinates ($q_1 = \varrho$, $q_2 = \varphi$, $q_3 = z$, $H_1 = 1$, $H_2 = \varrho$, $H_3 = 1$) we will have

$$d\mathbf{r} = \mathbf{e}_\varrho\, d\varrho + \mathbf{e}_\varphi \varrho\, d\varphi + \mathbf{e}_z\, dz,$$
$$\mathbf{a} = a_\varrho(\varrho, \varphi, z)\mathbf{e}_\varrho + a_\varphi(\varrho, \varphi, z)\mathbf{e}_\varphi + a_z(\varrho, \varphi, z)\mathbf{e}_z,$$

and by (24.155)

$$\int_L (\mathbf{a}, d\mathbf{r}) = \int_L a_\varrho \, d\varrho + \varrho a_\varphi \, d\varphi + a_z \, dz. \tag{24.156}$$

Similarly, for spherical coordinates ($q_1 = r$, $q_2 = \theta$, $q_3 = \varphi$, $H_1 = 1$, $H_2 = r$, $H_3 = r \sin\theta$) we will have

$$\mathbf{a} = a_r(r, \theta, \varphi)\mathbf{e}_r + a_\theta(r, \theta, \varphi)\mathbf{e}_\theta + a_\varphi(r, \theta, \varphi)\mathbf{e}_\varphi,$$
$$d\mathbf{r} = dr\,\mathbf{e}_r + r\,d\theta\,\mathbf{e}_\theta + r\sin\theta\,d\varphi\,\mathbf{e}_\varphi,$$

and by (24.155)

$$\int_L (\mathbf{a}, d\mathbf{r}) = \int_L a_r \, dr + r\,a_\theta \, d\theta + r\sin\theta a_\varphi \, d\varphi. \tag{24.157}$$

The circulation of a vector field $\mathbf{a}(M)$ in curvilinear coordinates q_1, q_2, q_3 is computed by formula (24.155) when the initial and terminal points of the curve L coincide (i.e., the curve L is closed), and in cylindrical and spherical coordinates by (24.156) and (24.157), respectively.

Example. Compute the circulation of a vector field given in cylindrical coordinates

$$\mathbf{a} = \varrho \sin\varphi \, \mathbf{e}_\varrho + \varrho z \, \mathbf{e}_\varphi + \varrho^3 \mathbf{e}_z$$

around a closed curve

$$L: \begin{cases} \varrho = \sin\varphi, \; 0 \leqslant \varphi \leqslant \pi, \\ z = 0. \end{cases}$$

◄ The coordinates of the vector are

$$a_\varrho = \varrho \sin\varphi, \qquad a_\varphi = \varrho z, \qquad a_z = \varrho^3.$$

Contour L is a closed curve lying in the plane $z = 0$ (Fig. 24.44). Substituting the coordinates of the vector into (24.156) gives the circulation

$$\oint_L \varrho \sin\varphi \, d\varrho + \varrho^2 z \, d\varphi + \varrho^3 dz.$$

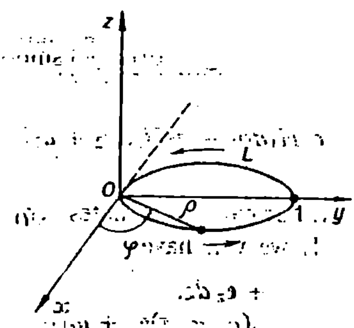


Fig. 24.44

On L we have

$$z = 0, \quad dz = 0, \quad \varrho = \sin\varphi, \quad d\varrho = \cos\varphi\, d\varphi, \quad 0 \leqslant \varphi \leqslant \pi.$$

The circulation will then be

$$\oint_L \varrho \sin\varphi\, d\varrho = \int_0^\pi \sin^2\varphi \cos\varphi\, d\varphi = \left.\frac{\sin^3\varphi}{3}\right|_0^\pi = 0. \blacktriangleright$$

Laplace operator in orthogonal coordinates. If $u = u(q_1, q_2, q_3)$ is a scalar function, then

$$\operatorname{grad} u = \frac{1}{H_1}\frac{\partial u}{\partial q_1}\mathbf{e}_1 + \frac{1}{H_2}\frac{\partial u}{\partial q_2}\mathbf{e}_2 + \frac{1}{H_3}\frac{\partial u}{\partial q_3}\mathbf{e}_3. \tag{*}$$

If

$$\mathbf{a} = a_1(q_1, q_2, q_3)\mathbf{e}_1 + a_2(q_1, q_2, q_3)\mathbf{e}_2 + a_3(q_1, q_2, q_3)\mathbf{e}_3,$$

then

$$\operatorname{div} \mathbf{a} = \frac{1}{H_1H_2H_3}\left[\frac{\partial}{\partial q_1}(a_1H_2H_3) + \frac{\partial}{\partial q_2}(a_2H_3H_1) + \frac{\partial}{\partial q_3}(a_3H_1H_2)\right]. \tag{**}$$

Using (*) and (**), we will obtain for the Laplacian the following expression:

$$\Delta u = \operatorname{div}\operatorname{grad} u = \frac{1}{H_1H_2H_3}\left[\frac{\partial}{\partial q_1}\left(\frac{H_2H_3}{H_1}\frac{\partial u}{\partial q_1}\right)\right.$$
$$\left. + \frac{\partial}{\partial q_2}\left(\frac{H_3H_1}{H_2}\frac{\partial u}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{H_1H_2}{H_3}\frac{\partial u}{\partial q_3}\right)\right].$$

Specifically, in cylindrical coordinates ($q_1 = \varrho$, $q_2 = \varphi$, $q_3 = z$, $H_1 = 1$, $H_2 = \varrho$, $H_3 = 1$) we will get

$$\Delta u = \frac{1}{\varrho}\left[\frac{\partial}{\partial\varrho}\left(\varrho\frac{\partial u}{\partial\varrho}\right) + \frac{\partial}{\partial\varphi}\left(\frac{1}{\varrho}\frac{\partial u}{\partial\varphi}\right) + \frac{\partial}{\partial z}\left(\varrho\frac{\partial u}{\partial z}\right)\right]$$
$$= \frac{1}{\varrho}\frac{\partial}{\partial\varrho}\left(\varrho\frac{\partial u}{\partial\varrho}\right) + \frac{1}{\varrho^2}\frac{\partial^2 u}{\partial\varphi^2} + \frac{\partial^2 u}{\partial z^2}.$$

In spherical coordinates ($q_1 = r$, $q_2 = \theta$, $q_3 = \varphi$, $H_1 = 1$, $H_2 = r$, $H_3 = r\sin\theta$) we will have

$$\Delta u = \frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial u}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) + \frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\theta}\frac{\partial u}{\partial\varphi}\right)\right]$$
$$= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\varphi^2}.$$

Example. Find all solutions of the Laplace equation $\Delta u = 0$ that only depend on the distance r.

◄ Since the desired solutions must only depend on the distance r of the point M from the origin of coordinates, i.e., $u = u(r)$, then the Laplace equation $\Delta u = 0$ in spherical coordinates will have the form

$$\Delta u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) = 0.$$

Hence

$$r^2\frac{\partial u}{\partial r} = \tilde{C}_1,$$

so that

$$u(r) = \frac{C_1}{r} + C_2,$$

where C_1 and C_2 are constants. ►

Exercises

Find the derivatives of the scalar field $u(x, y, z)$ at a point $M_0(x_0, y_0, z_0)$ in the direction of $M_1(x_1, y_1, z_1)$:

1. $u = x\sqrt{y} + y\sqrt{z}$, $M_0(2, 4, 4)$, $M_1(6, -4, 8)$.
2. $u = 4\ln(x^2 + 3) - 8xyz$, $M_0(1, 1, 1)$, $M_1(3, -3, 5)$.
3. $u = x^3 + \sqrt{y^2 + z^2}$, $M_0(1, -3, 4)$, $M_1(1, -2, 3)$.

Find the derivative of the scalar field $u(x, y, z)$ at a point $M_0(x_0, y_0, z_0)$ along the normal to a surface S that forms an acute angle with the positive z-axis:

4. $u = \tan^{-1}\frac{y}{x} + xz$, $M_0(2, 2, -1)$, S: $x^2 + y^2 - 2z = 10$.
5. $u = \sqrt{xy} - \sqrt{4 - z^2}$, $M_0(1, 1, 0)$, S: $x^2 - y^2 = z$.
6. Find the derivative of the scalar field $u = 2xy + y^2$ at a point $M_0(\sqrt{2}, 1)$ of the ellipse $x^2/4 + y^2/2 = 1$ along the external normal to the ellipse at that point.
7. Find the derivative of the scalar field $u = z\ln(x^2 + y^2 - z)$ at the point $M_0(1, -\sqrt{3}, 3)$ in the direction of the circle $x = 2\cos t$, $y = 2\sin t$, $z = 3$.
8. Find the angle between the gradients of the function $u = \tan^{-1}(x/y)$ at the points $M_1(1, 1)$ and $M_2(-1, -1)$.
9. Find the derivative of the plane field $u = x^3 + xy + 3y^4$ at the point $M(-2, 1, 0)$ along a line lying in the xy-plane and inclined at an angle $\pi/6$ to the x-axis.

Find the vector lines of the following vector fields:

10. $\mathbf{a} = x\mathbf{i} + 4y\mathbf{j}$.
11. $\mathbf{a} = (z - y)\mathbf{i} + (x - z)\mathbf{j} + (y - x)\mathbf{k}$.
12. $\mathbf{a} = 2z\mathbf{j} + 3y\mathbf{k}$.
13. Find the vector line of the field $\mathbf{a} = x^2\mathbf{i} - y^3\mathbf{j} + z^2\mathbf{k}$ that passes through the point $M(1/2, -1/2, 1)$.
14. Find the vector line of the field $\mathbf{a} = -y\mathbf{i} + x\mathbf{j}$ that passes through the point $M(3, 4, -1)$.
15. Find the flux of the vector field $\mathbf{a} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ through the upper side of the circle cut out by the cones $x^2 + y^2 = z^2$ in the plane $z = h (h > 0)$.
16. Find the flux of the vector field $\mathbf{a} = (x - 2z)\mathbf{i} + (x + 3y + z)\mathbf{j} + (5x + y)\mathbf{k}$ through the triangle ABC with vertices at the points $A(1, 0, 0)$, $B(0, 1, 0)$, $C(0, 0, 1)$. The normal forms an acute angle with the z-axis.
17. Find the flux of the vector field $\mathbf{a} = x\mathbf{i} + z\mathbf{k}$ through the side surface of the circular cylinder $x^2 + y^2 = R^2$ bounded by the planes $z = 0$, $z = h$ $(h > 0)$.
18. Find the flux of the vector field $\mathbf{a} = yz\mathbf{i} - x\mathbf{j} - y\mathbf{k}$ through the total surface of the cone $x^2 + y^2 = z^2$ bounded by the plane $z = 1$ $(0 \leqslant z \leqslant 1)$.

By introducing curvilinear coordinates on a surface determine the fluxes of vector $\mathbf{a}$ through a surface S:

19. $\mathbf{a} = x^3\mathbf{i} - y^3\mathbf{j} + xz^3\mathbf{k}$, S is the external side of the cylindrical surface $x^2 + y^2 = 9$ bounded by the sphere $x^2 + y^2 + z^2 = 25$.
20. $\mathbf{a} = x^3\mathbf{i} - y^3\mathbf{j} + z\mathbf{k}$, S is the external side of the sphere $x^2 + y^2 + z^2 = 1$ cut out by the conical surface $x^2 + y^2 = z^2$ (where $z \geqslant \sqrt{x^2 + y^2}$).

Find the flux of a vector field $\mathbf{a}$ through a closed surface S (external normal). Test the result using the Ostrogradsky formula.

21. $\mathbf{a} = yz\mathbf{j}$, S: $\{x^2 + y^2 = 1 - z,\ z = 0,\ y \geqslant 0\}$.
22. $\mathbf{a} = y^2\mathbf{j} + 2z\mathbf{k}$, S: $\{x^2 + y^2 = 4,\ x^2 + y^2 = z,\ z = 0\}$.
23. $\mathbf{a} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$, S: $\{x^2 + y^2 + z^2 = R^2,\ z = 0\ (z > 0)\}$.
24. $\mathbf{a} = yz\mathbf{i} - x\mathbf{j} + y\mathbf{k}$, S: $\{x^2 + z^2 = y^2,\ y = 1\ (0 \leqslant y \leqslant 1)\}$.

By adequately closing open surfaces and using the Ostrogradsky-Gauss theorem, find the fluxes of vector fields through surfaces (we take the external normal to the closed surface).

25. $\mathbf{a} = (1 - 2x)\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, S: $\{x^2 + y^2 = z^2\ (0 \leqslant z \leqslant 4)\}$.
26. $\mathbf{a} = z^2\mathbf{i} + xz\mathbf{j} + y\mathbf{k}$, S: $\{x^2 + y^2 = 4 - z\ (z \geqslant 0)\}$.
27. $\mathbf{a} = y\mathbf{i} - 2x\mathbf{j} - z\mathbf{k}$, S: $\{x^2 + y^2 + z^2 = 4\ (z \geqslant 0)\}$.

Find the work done by a force $\mathbf{F}$ in moving along a line L from point M to point N.

28. $\mathbf{F} = (x^2 - 2y)\mathbf{i} + (y^2 - 2x)\mathbf{j}$, L: {segment MN, $M(-4, 0)$, $N(0, 2)$}.

29. $\mathbf{F} = (x + y)\mathbf{i} + 2x\mathbf{j}$, L: $\{x^2 + y^2 = 4\ (y \geqslant 0),\ M(2, 0),\ N(-2, 0)\}$.

30. $\mathbf{F} = (x + y)\mathbf{i} + (x - y)\mathbf{j}$, L: $\left\{x^2 + \frac{y^2}{9} = 1\ (x \geqslant 0,\ y \geqslant 0),\ M(1, 0),\ N(0, 3)\right\}$.

31. $\mathbf{F} = -y\mathbf{i} + x\mathbf{j}$, L: $\{y = x^3,\ M(0, 0),\ N(2, 8)\}$.

Find the circulation of a vector field $\mathbf{a}$ around a closed contour L (in the direction where t grows).

32. $\mathbf{a} = x\mathbf{i} - z^2\mathbf{j} + y\mathbf{k}$, L: $\{x = 2\cos t,\ y = 3\sin t,\ z = 4\cos t - 3\sin t - 3\}$.

33. $\mathbf{a} = (y - z)\mathbf{i} + (z - x)\mathbf{j} + (x - y)\mathbf{k}$, L: $\{x = 4\cos t,\ y = 4\sin t,\ z = 1 - \cos t\}$.

34. $\mathbf{a} = -x^2y^3\mathbf{i} - 2\mathbf{j} + xz\mathbf{k}$, L: $\{x = \sqrt{2}\cos t,\ y = \sqrt{2}\sin t,\ z = 1\}$.

Find the circulation of a vector field $\mathbf{a}$ around a closed contour L. Test the results using Stokes' theorem.

35. $\mathbf{a} = 2y\mathbf{i} - z\mathbf{j} + x\mathbf{k}$, L: $\{x^2 + y^2 = 1,\ x + y + z = 4\}$.

36. $\mathbf{a} = xyz\mathbf{i} + (x + y + z)\mathbf{j} - x^2y^2\mathbf{k}$, L: $\{x + y = a,\ x - y = a,\ x + y = -a,\ x - y = -a\}$.

37. $\mathbf{a} = 2xz\mathbf{i} - y\mathbf{j} + z\mathbf{k}$, where L is the line of intersection of the plane $x + y + 2z = 2$ with the coordinate planes $x = 0$, $y = 0$, $z = 0$.

38. Find the divergence of the vector field $\mathbf{a} = [\mathbf{c}, \mathbf{r}]$, where $\mathbf{c}$ is a constant vector and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

39. For which function $\psi(z)$ the divergence of the vector field $\mathbf{a} = xz\mathbf{i} + y\mathbf{j} + \psi(z)\mathbf{k}$ will be equal to z?

40. Find $\operatorname{div}(r^4\mathbf{r})$, where $r = \sqrt{x^2 + y^2 + z^2}$, and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

41. Find the function $\psi(r)$ for which $\operatorname{div}(\psi(r)\mathbf{r}) = 2\psi(r)$.

42. Find the function $f(x, z)$ such that the curl of the vector field $\mathbf{a} = yz\mathbf{i} + f(x, z)\mathbf{j} + xy\mathbf{k}$ would coincide with the vector $\mathbf{k} - \mathbf{i}$.

Find the curls of the vectors:

43. $\mathbf{a} = y^2z\mathbf{i} + xz^2\mathbf{j} + x^2y\mathbf{k}$.

44. $\mathbf{a} = y\mathbf{i} - x\mathbf{j} + z\mathbf{k}$.

45. $\mathbf{a} = 2xz\mathbf{i} - y\mathbf{j} + z\mathbf{k}$.

Prove that the following vector fields are potential fields and find their potentials:

46. $\mathbf{a} = \dfrac{x\mathbf{i} + y\mathbf{j}}{\sqrt{x^2 + y^2 + 4}}$.

47. $\mathbf{a} = (yz + 1)\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}$.

48. $\mathbf{a} = \dfrac{\mathbf{i} + \mathbf{j} + \mathbf{k}}{x + y + z}$.

Answers

1. $-\dfrac{\sqrt{6}}{3}$. 2. -2. 3. $-0.7\sqrt{2}$. 4. $\dfrac{3}{3}$. 5. 0. 6. $\dfrac{2\sqrt{3}(\sqrt{2}+3)}{3}$. 7. 0. 8. π. 9. $\dfrac{13\sqrt{3}}{2}+5$. 10. $y = C_1x^4$, $z = C_2$. 11. $x + y + z = C_1$, $x^2 + y^2 + z^2 = C_2^2$. 12. $3y^2 - 2z^2 = C_1$, $x = C_2$. 13. $\dfrac{1}{x} - \dfrac{1}{z} = 1$, $\dfrac{1}{x} + \dfrac{1}{2y^2} = 4$. 14. $x^2 + y^2 = 25$, $z = -1$. 15. πh^3. 16. $\dfrac{5}{3}$. 17. πhR^2. 18. 0. 19. 0. 20. $\dfrac{2\pi}{3}\left(1 - \dfrac{\sqrt{2}}{4}\right)$. 21. $\dfrac{\pi}{12}$. 22. 16π. 23. $\dfrac{\pi R^4}{2}$. 24. 0. 25. -64π. 26. 0. 27. $-\dfrac{16\pi}{3}$. 28. 24. 29. 2π. 30. -5. 31. 8. 32. 60π. 33. -40π. 34. π. 35. -2π. 36. $2a^2$. 37. $\dfrac{4}{3}$. 38. 0. 39. $\psi(z) = C - z$, $C = \text{const}$. 40. $7r^4$. 41. $\psi(r) = \dfrac{C}{r}$, $C = \text{const}$. 42. $f(x, z) = (1 + z)x + z + C$, $C = \text{const}$. 43. $(x^2 - 2xz)\mathbf{i} + (y^2 - 2xy)\mathbf{j} + (z^2 - 2yz)\mathbf{k}$. 44. $-2\mathbf{k}$. 45. $2x\mathbf{j}$. 46. $\sqrt{x^2 + y^2 + 4} + C$, $C = \text{const}$. 47. $x(1 + yz) + C$, $C = \text{const}$. 48. $\ln|x + y + z| + C$, $C = \text{const}$.

Chapter 25

Integrals Depending on Parameter

25.1 Proper Integrals Depending on Parameter

Concept of an integral depending on parameter. Continuity of the integral. Let $f(x, y)$ be a function of two variables defined in the rectangle (Fig. 25.1)

$$\Pi: \{a \leqslant x \leqslant b, \quad c \leqslant y \leqslant d\}.$$

Suppose that for any fixed value of $y \in [c, d]$ there exists the integral

$$\int_a^b f(x, y)\,dx,$$

which is clearly a function of y, i.e.,

$$I(y) = \int_a^b f(x, y)\,dx, \qquad y \in [c, d]. \tag{25.1}$$

Integral (25.1) is called a *proper integral depending on the parameter y.*

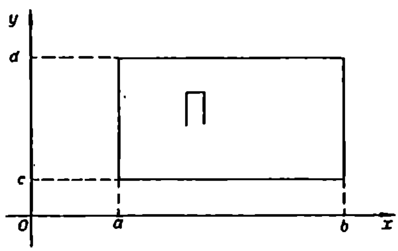


Fig. 25.1

The continuity of the integral depending on a parameter follows from the following theorem:

Theorem 25.1. *If a function $f(x, y)$ is continuous in the rectangle Π, then the function $I(y)$ given by* (25.1) *is continuous on the segment* $[c, d]$.

◄ It follows from (25.1) that the increment $\Delta I = I(y + \Delta y) - I(y)$ of the function $I(y)$ corresponding to the increment of Δy can be estimated to be

$$|I(y + \Delta y) - I(y)| = \left| \int_a^b f(x, y + \Delta y)\,dx - \int_a^b f(x, y)\,dx \right|$$
$$= \left| \int_a^b [f(x, y + \Delta y) - f(x, y)]\,dx \right|$$
$$\leqslant \int_a^b |f(x, y + \Delta y) - f(x, y)|\,dx. \tag{25.2}$$

By the conditions of the theorem $f(x, y)$ is continuous in the closed rectangle Π, and hence $f(x, y)$ is uniformly continuous in that rectangle. Accordingly, for any $\varepsilon > 0$ we can find $\delta > 0$ such that for all x from $[a, b]$ and all y and $y + \Delta y$ from $[c, d]$ such that $|\Delta y| < \delta$ we will have $|f(x, y + \Delta y) - f(x, y)| < \varepsilon/|b - a|$. It follows from this and the estimate (25.2) that for $|\Delta y| < \delta$

$$|I(y + \Delta y) - I(y)| \leqslant \frac{\varepsilon}{b - a} \int_a^b dx = \varepsilon.$$

This implies that the function $I(y)$ is continuous at each point of the segment $[c, d]$. ►

Corollary. If $f(x, y)$ is continuous in Π, then

$$\lim_{y \to y_0} \int_a^b f(x, y)\,dx = \int_a^b \lim_{y \to y_0} f(x, y)\,dx = \int_a^b f(x, y_0)\,dx, \tag{25.3}$$

where y_0 is any fixed number belonging to $[c, d]$.

◄ In fact, since $I(y)$ is continuous on $[c, d]$,

$$\lim_{y \to y_0} I(y) = I(\lim_{y \to y_0} y) = I(y_0).$$

These equalities are equivalent to (25.3). ►

Example. Find

$$\lim_{y \to 0} \int_1^2 (2x - 1)\cos(xy)\,dx.$$

◄ Since the function $f(x, y) = (2x - 1)\cos(xy)$ is continuous in any Π: $\{1 \leqslant x \leqslant 2,\ c \leqslant y \leqslant d\}$, where $c < 0 < d$, by (25.3) we have

$$\lim_{y \to 0} \int_1^2 (2x - 1)\cos(xy)\,dx = \int_1^2 (2x - 1)\,dx = 2. \text{►}$$

Differentiation of an integral with respect to a parameter.

Theorem 25.2. *If a function $f(x, y)$ and its partial derivative $\partial f(x, y)/\partial y$ are continuous in a rectangle Π, then there holds the Leibniz formula of differentiation with respect to the parameter in the integrand*

$$I'(y) = \int_a^b \frac{\partial f(x, y)}{\partial y}\, dx. \tag{25.4}$$

◄ Assuming that $y + \Delta y \in [c, d]$, we set up the difference relation

$$\frac{I(y + \Delta y) - I(y)}{\Delta y} = \int_a^b \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}\, dx.$$

Passing to the limit as $\Delta y \to 0$ and using the continuity property of the partial derivative $\partial f(x, y)/\partial y$, we will get by (25.3)

$$I'(y) = \lim_{\Delta y \to 0} \frac{I(y + \Delta y) - I(y)}{\Delta y}$$

$$= \int_a^b \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}\, dx = \int_a^b \frac{\partial f(x, y)}{\partial y}\, dx. \text{ ►}$$

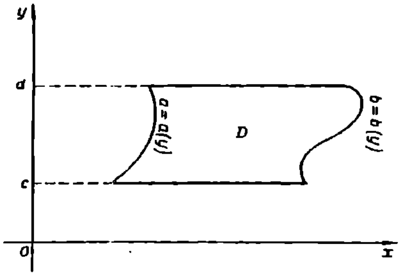


Fig. 25.2

Remark. Let the limits of integration be dependent on the parameter y. Then

$$F(y) = \int_{a(y)}^{b(y)} f(x, y)\, dx, \tag{25.5}$$

where $a(y) \leqslant x \leqslant b(y)$ and the functions $a(y)$ and $b(y)$are differentiable on the interval $c \leqslant y \leqslant d$. If $f(x, y)$ and $f'_y(x, y)$ are continuous in a domain D: $\{a(y) \leqslant x \leqslant b(y), c \leqslant y \leqslant d\}$ (Fig. 25.2), we find that $f(y)$ is differentia-

ble on $[c, d]$ and

$$F'(y) = \int_{a(y)}^{b(y)} \frac{\partial f(x, y)}{\partial y} dx + f[b(y), y]b'(y) - f[a(y), y]a'y. \quad (25.6)$$

◄ Formula (25.6) is proved by differentiating a composite function. Since $f(y) = f(y, a(y), b(y))$, the total derivative will be

$$F'(y) = \frac{\partial F}{\partial y} + \frac{\partial F}{\partial a}\frac{da}{dy} + \frac{\partial F}{\partial b}\frac{db}{dy}, \quad (25.7)$$

where

$$\frac{\partial F}{\partial y} = \int_{a(y)}^{b(y)} \frac{\partial f(x, y)}{\partial y} dx,$$

$$\frac{dF}{\partial b} = \frac{\partial}{\partial b}\int_{a(y)}^{b(y)} f(x, y)\, dx = f[b(y), y],$$

$$\frac{\partial F}{\partial a} = \frac{\partial}{\partial a}\int_{a(y)}^{b(y)} f(x, y)\, dx = -f[a(y), y].$$

Substituting $\partial F/\partial y$, $\partial F/\partial b$ and $\partial F/\partial a$ into (25.7) gives (25.6). ►

Examples. (1) Differentiating with respect to the parameter, take the integral

$$I(a) = \int_0^{\pi/2} \ln\frac{1 + a\cos x}{1 - a\cos x}\,\frac{dx}{\cos x},$$

where $|a| < 1$.

◄ The function

$$f(x, a) = \begin{cases} \dfrac{1}{\cos x}\ln\dfrac{1 + a\cos x}{1 - a\cos x} & \text{at } x \neq \dfrac{\pi}{2}, \\ 2a & \text{at } x = \dfrac{\pi}{2}, \end{cases}$$

and its derivative with respect to the parameter

$$f'_a(x, a) = \frac{2}{1 - a^2\cos^2 x}$$

are continuous in the rectangle

$$\Pi: \left\{0 \leqslant x \leqslant \frac{\pi}{2},\quad |a| \leqslant 1 - \varepsilon < 1\right\}.$$

Therefore, we can here apply Theorem 25.2 for $|a| \leqslant 1 - \varepsilon < 1$. We thus obtain

$$I'(a) = \int_0^{\pi/2} \frac{\partial}{\partial a}\left(\frac{1}{\cos x}\ln\frac{1 + a\cos x}{1 - a\cos x}\right)dx = 2\int_0^{\pi/2} \frac{dx}{1 - a^2\cos^2 x}.$$

We put $\tan x = t$, then $dx = dt/(1 + t^2)$, $\cos^2 x = 1/(1 + t^2)$, and integrate with respect to t from 0 to $+\infty$ so that

$$I'(a) = 2\int_0^{+\infty} \frac{dt}{1 + t^2 - a^2} = 2\int_0^{+\infty} \frac{dt}{(\sqrt{1 - a^2})^2 + t^2}$$

$$= \frac{2}{\sqrt{1 - a^2}}\tan^{-1}\frac{t}{\sqrt{1 - a^2}}\Bigg|_0^{+\infty} = \frac{\pi}{\sqrt{1 - a^2}}.$$

Hence $I(a) = \pi\sin^{-1}a + C$.

When ε tends to zero, we see that this result holds for $|a| < 1$. Since from the original integral we get $I(0) = 0$, then $C = 0$, and so $I(a) = \pi\sin^{-1}a$. ►

(2) Find $F'(y)$ for $F(y) = \int_y^{y^2} e^{-x^2 y}\,dx$.

◄ Here $f(x, y) = e^{-x^2 y}$, $a(y) = y$, $b(y) = y^2$. Using (25.6) gives

$$F'(y) = -\int_y^{y^2} x^2 e^{-x^2 y}\,dx + e^{-y^5}2y - e^{-y^3}.$$ ►

Integration of the integral depending on a parameter. We will need the following theorem.

Theorem 25.3. *If a function $f(x, y)$ is continuous in a rectangle Π then the function $I(y) = \int_a^b f(x, y)\,dx$ is integrable on the interval $[c, d]$ and*

$$\int_c^d I(y)\,dy = \int_c^d \left[\int_a^b f(x, y)\,dx\right]dy = \int_a^b \left[\int_c^d f(x, y)\,dy\right]dx. \tag{25.8}$$

In other words, if $f(x, y)$ is continuous in Π, then the integral depending on the parameter can be integrated with respect to the parameter in the integrand.

◄ According to Theorem 25.1 the function $I(y)$ is continuous on $[c, d]$ and is therefore integrable on it.

Formula (25.8) is valid because the repeated integrals in it are equal (see below), since

$$\int_c^d \left[\int_a^b f(x, y)\, dx \right] dy = \int_a^b \left[\int_c^d f(x, y)\, dy \right] dx = \iint_\Pi f(x, y)\, dx\, dy. \blacktriangleright$$

Example. Integrate with respect to the parameter y the integral

$$I(y) = \int_a^b y^x\, dx \qquad (0 < a < b)$$

from 0 to 1.

◄ Since $f(x, y) = y^x$ is continuous in the rectangle

$$\Pi: \{a \leqslant x \leqslant b, \quad 0 \leqslant y \leqslant 1\} \qquad (a > 0),$$

we can apply Theorem 25.3. We have

$$\int_0^1 I(y)\, dy = \int_0^1 dy \int_a^b y^x\, dx = \int_a^b dx \int_0^1 y^x\, dy$$

$$= \int_a^b \left(\frac{y^{x+1}}{x+1} \Bigg|_{y=0}^{1} \right) dx = \int_a^b \frac{dx}{x+1} = \ln \frac{b+1}{a+1}. \blacktriangleright$$

25.2 Improper Integrals Depending on Parameter

Concept of the improper integral of the first kind depending on a parameter. Let a function $f(x, y)$ be defined in the domain (Fig. 25.3)

$$\Pi_\infty: \{a \leqslant x < +\infty, \qquad c \leqslant y \leqslant d\}.$$

Suppose that for each fixed $y \in [c, d]$ there exists the improper integral $\int_a^{+\infty} f(x, y)\, dx$ that is a function of y. Then the function

$$I(y) = \int_a^{+\infty} f(x, y)\, dx, \qquad y \in [c, d], \tag{25.9}$$

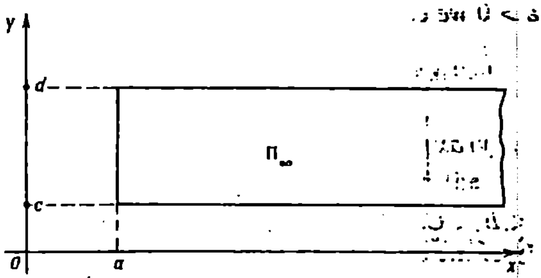


Fig. 25.3

is called the *importer integral of the first kind depending on the parameter y*. The interval $[c, d]$ here can also be infinite.

Definition. The improper integral (25.9) is called *convergent* at the point $y \in [c, d]$, if there exists the finite limit

$$\lim_{B \to +\infty} \int_a^B f(x, y)\, dx = I(y),$$

i.e., if for any $\varepsilon > 0$ there exists a number B_0, such that for all $B \geqslant B_0$ we have

$$\left| I(y) - \int_a^B f(x, y)\, dx \right| < \varepsilon.$$

If the improper integral (25.9) converges at each point y of the interval $[c, d]$, then it is said to be *convergent* on that interval. Integral (25.9) is said to be *absolutely convergent* on $[c, d]$ if the integral

$$\int_a^\infty |f(x, y)|\, dx$$

is convergent.

Uniform convergence of the improper integral. Cauchy criterion.

Definition. The improper integral (25.9) is said to be *uniformly convergent in the parameter y on the interval* $[c, d]$, if it is convergent on the interval and if for any $\varepsilon > 0$, we can indicate $A \geqslant a$, dependent only on ε, such that for all $B > A$ and all y belonging to $[c, d]$ holds the inequality

$$\left| \int_B^\infty f(x, y)\, dx \right| < \varepsilon, \qquad B > A. \tag{25.10}$$

There exists the following Cauchy criterion for the uniform convergence of improper integrals depending on a parameter.

Theorem 25.4. *For the improper integral* (25.9) *to be uniformly convergent in the parameter y on the interval* $[c, d]$ *it is necessary and sufficient that for any* $\varepsilon > 0$ *we could indicate a number* $A \geqslant a$, *depending only on* ε, *such that for any* $B > A$ *and* $C > A$ *and for all y from the interval* $[c, d]$ *holds the inequality*

$$\left| \int_B^C f(x, y)\, dx \right| < \varepsilon, \tag{25.11}$$

where $a \leqslant A < B < C$.

The validity of the criterion follows directly from the definition of uniform convergence.

We now prove the sufficient test for the uniform convergence of improper integrals depending on a parameter.

Theorem 25.5 (Weierstrass test). *Let a function $f(x, y)$ be defined in a half-strip Π_∞ and be integrable with respect to x for each $y \in [c, d]$ on any interval $[a, A]$. Suppose also that for all the points on the half-strip holds the inequality*

$$|f(x, y)| \leqslant g(x). \tag{25.12}$$

Then from the fact that the integral $\int_a^\infty g(x)\,dx$ converges, it follows that the improper integral $I(y) = \int_a^\infty f(x, y)\,dx$ depending on the parameter y is uniformly convergent in y on $[c, d]$.

◄ According to the Cauchy criterion for the convergence of the integral of a function $g(x)$ for any $\varepsilon > 0$, we can find a number $A \geqslant a$ such that for all $C > B \geqslant A$ holds the inequality

$$\int_B^C g(x)\,dx < \varepsilon.$$

Using (25.12), we find

$$\left|\int_B^C f(x, y)\,dx\right| \leqslant \int_B^C g(x)\,dx < \varepsilon$$

for all y on $[c, d]$, which means that the Cauchy criterion for the uniform convergence of the integral

$$I(y) = \int_a^\infty f(x, y)\,dx$$

is valid. ►

Example. Examine for uniform convergence the improper integral

$$I(s) = \int_0^\infty e^{-x}\sin(sx)\,dx, \tag{25.13}$$

where s is a parameter, $s \in [\alpha, \beta]$.

◄ For any $s \in [\alpha, \beta]$, where α and β are arbitrary real numbers, we have

$$|e^{-x}\sin(sx)| \leqslant e^{-x}$$

and the integral $\int_0^\infty e^{-x}\,dx = 1$ is convergent. By the Weierstrass test the integral (25.13) converges uniformly for all $s \in [\alpha, \beta]$. We can show that $I(s) = s/(1 + s^2)$. ►

Properties of uniformly convergent improper integrals depending on a parameter. Uniformly convergent improper integrals depending on a parameter have the following properties. We provide them without proof.

(1) *Continuity.* If a function $f(x, y)$ is continuous in a domain Π_∞ and the integral

$$I(y) = \int_a^\infty f(x, y)\,dx \tag{25.14}$$

converges uniformly in y on $[c, d]$, then the function $I(y)$ is continuous on $[c, d]$.

(2) *Integrability.* If a function $f(x, y)$ is continuous in a domain Π_∞ and the integral (25.14) converges uniformly in y on $[c, d]$, then

$$\int_c^d I(y)\,dy = \int_c^d \left[\int_a^\infty f(x, y)\,dx\right] dy = \int_a^\infty \left[\int_c^d f(x, y)\,dy\right] dx. \tag{25.15}$$

(3) *Differentiability.* Let a function $f(x, y)$ and its partial derivative $\partial f(x, y)/\partial y$ be continuous in Π_∞. Suppose further that (25.14) converges, and the integral

$$\int_a^\infty \frac{\partial f(x, y)}{\partial y}\,dx$$

converges uniformly in y on $[c, d]$. Then

$$I'(y) = \int_a^\infty \frac{\partial f(x, y)}{\partial y}\,dx. \tag{25.16}$$

Examples. (1) Take the integral

$$K(s) = \int_0^\infty x e^{-x} \cos(sx)\,dx, \tag{25.17}$$

which depends on the parameter s.

◄ We have already established that the integral (25.13) converges uniformly in parameter s on any interval $[\alpha, \beta]$.

◄ We have already established that the integral (25.13) converges uniformly in parameter s on any interval $[\alpha, \beta]$.

Indeed, for any s we first have

$$|x e^{-x} \cos(sx)| \leqslant x e^{-x},$$

second

$$\int_0^\infty x \cdot e^{-x}\,dx = 1.$$

So, using the Weierstrass test, we conclude that integral (25.17) converges uniformly.

Denoting by $f(x, s)$ the integrand of (25.13), i.e.,

$$f(x, s) = e^{-x}\sin(sx),$$

we notice that $\partial f/\partial s = x e^{-x}\cos(sx)$ is the integrand of (25.17), which, as we have established, converges uniformly in the parameter s on any $[\alpha, \beta]$.

We now use the differentiability property of the improper integral in parameter. We have $K(s) = I'(s)$. But $I(s) = s/(1 + s^2)$, therefore

$$K(s) = \left(\frac{s}{1 + s^2}\right)' = \frac{1 - s^2}{(1 + s^2)^2}.$$

Thus

$$\int_0^\infty x e^{-x}\cos(sx)\,dx = \frac{1 - s^2}{(1 + s^2)^2}. \blacktriangleright$$

(2) Integrating $\int\limits_0^\infty e^{-xy}\,dx = \frac{1}{y}$ with respect to y ($y > 0$), find the integral

$$\int_0^\infty \frac{e^{-ax} - e^{-bx}}{x}\,dx, \qquad 0 < a < b.$$

We assume that at $x = 0$

$$\lim_{x\to 0} \frac{e^{-ax} - e^{-bx}}{x} = b - a.$$

◄ We show at first that the improper integral

$$I(y) = \int_0^\infty e^{-xy}\,dx, \qquad 0 < a \leqslant y \leqslant b$$

depending on the parameter y converges uniformly on $[a, b]$.

This follows from the Weierstrass test, since

$$|f(x, y)| = e^{-xy} \leqslant e^{-ax} \qquad (a > 0)$$

and

$$\int_0^\infty e^{-ax}\,dx = \frac{1}{a}$$

(the integral converges). We now integrate $I(y) = \int\limits_0^\infty e^{-xy}\,dx$ with respect to the parameter y from a to b:

$$\int\limits_a^b \left(\int\limits_0^\infty e^{-xy}\,dx \right) dy = \int\limits_0^\infty \left(\int\limits_a^b e^{-xy}\,dy \right) dx$$

$$= \int\limits_0^\infty \left(-\frac{e^{-xy}}{x} \right) \Bigg|_{y=a}^{b} dx = \int\limits_0^\infty \frac{e^{-ax} - e^{-bx}}{x}\,dx.$$

But $\int\limits_0^\infty e^{-xy}\,dx = 1/y$, therefore

$$\int\limits_a^b \frac{dy}{y} = \ln y \Big|_a^b = \ln \frac{b}{a},$$

so that

$$\int\limits_0^\infty \frac{e^{-ax} - e^{-bx}}{x}\,dx = \ln \frac{b}{a}. \blacktriangleright$$

Remark. We have so far dealt with improper integrals of the form

$$I(y) = \int\limits_a^\infty f(x, y)\,dx, \qquad y \in [c, d].$$

These are improper integrals of the first kind dependent on the parameter y. The *improper integral of the second kind depending on the parameter* y is an integral of the form

$$F(y) = \int\limits_a^b f(x, y)\,dx, \tag{25.18}$$

where $\lim\limits_{x \to \xi} f(x, y) = \pm\infty$, $a \leqslant \xi \leqslant b$, $c \leqslant y \leqslant d$. The interval (a, b) may well be infinite.

The theory of improper integrals of the second kind depending on a parameter is similar to the theory of the first-kind improper integrals considered above.

25.3 Euler Integrals. Gamma Function. Beta Function

The integral

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t}\, dt \tag{25.19}$$

which is a function of the parameter x, is called the *gamma function.*

Turning to the range of the gamma function $\Gamma(x)$, we notice two types of singularities:

(1) integrated along the ray $0 \leqslant t < +\infty$;

(2) for $x < 1$ the point $t = 0$ is a singularity of the integrand (which becomes infinite).

To separate these singularities we represent $\Gamma(x)$ as a sum of two integrals

$$\Gamma(x) = \int_0^1 e^{-t} t^{x-1}\, dt + \int_1^{+\infty} e^{-t} t^{x-1}\, dt = I_1(x) + I_2(x).$$

Since $|e^{-t} t^{x-1}| \leqslant t^{x-1}$ for $t > 0$, the integral $I_1(x)$ converges for $x > 0$ (by the comparison test). The integral $I_2(x)$ converges for any x. Really, if we take $\lambda > 1$, we will have for any x

$$\lim_{t \to +\infty} \frac{e^{-t} t^{x-1}}{1/t^{\lambda}} = \lim_{t \to +\infty} \frac{t^{\lambda + x - 1}}{e^{t}} = 0.$$

But the integral $\int_1^{\infty} \frac{dt}{t^{\lambda}}$ converges for $\lambda > 1$, and so does the integral $\int_1^{\infty} e^{-t} t^{x-1}\, dt$ for any x. Hence $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\, dt$ converges for $x > 0$.

We have proved that the gamma function $\Gamma(x)$ is defined on the ray $x > 0$.

We now show that integral (25.19) converges uniformly in x on any interval $[c, d]$ where $0 < c < d < +\infty$.

Indeed let $c \leqslant x \leqslant d$. Then for $0 \leqslant t \leqslant 1$

$$\int_0^1 t^{x-1} e^{-t}\, dt \leqslant \int_0^1 t^{c-1} e^{-t}\, dt, \tag{25.20}$$

and for $t \geqslant 1$

$$\int_1^{\infty} t^{x-1} e^{-t}\, dt \leqslant \int_1^{\infty} t^{d-1} e^{-t}\, dt. \tag{25.21}$$

The integrals on the right-hand sides of (25.20) and (25.21) converge, and so do (by the Weierstrass test) the integrals on the left-hand sides of these relations.

Accordingly, by

$$\Gamma(x) = \int_0^1 t^{x-1} e^{-t}\,dt + \int_1^{\infty} e^{-t} t^{x-1}\,dt, \tag{25.19'}$$

we find that $\Gamma(x)$ is uniformly convergent on any $[c, d]$, where $0 < c < d$.

Since $\Gamma(x)$ is uniformly convergent, it is continuous for $x > 0$.

Some properties of the gamma function. (1) $\Gamma(x) > 0$ for $x > 0$, i.e., the gamma function has no zeros for $x >$, (2) $\Gamma(x + 1) = \int_0^{\infty} t^x e^{-t}\,dt =$

$$\left| \begin{matrix} u = t^x, du = xt^{x-1}\,dt \\ dv = e^{-t}\,dt, \; v = -e^{-t} \end{matrix} \right| = -t^x e^{-t}\Big|_{t=0}^{+\infty} + x\int_0^{\infty} t^{x-1} e^{-t}\,dt = x\Gamma(x).$$

Therefore, for any $x > 0$ we have the following formula:

$$\Gamma(x + 1) = x\Gamma(x). \tag{25.22}$$

(3) For integral $x = n$ there holds

$$\Gamma(n + 1) = n! \tag{25.23}$$

◄ Indeed,

$$\Gamma(1) = \int_0^{\infty} e^{-t}\,dt = 1.$$

Therefore, using (25.22), we will obtain

$$\begin{aligned}\Gamma(n + 1) &= n\Gamma(n) = n(n - 1)\Gamma(n - 1) \\ &= \ldots = n(n - 1)\ldots 3 \times 2 \times 1 \times \Gamma(1)\end{aligned}$$

i.e., $\Gamma(n + 1) = n!$ ►

At $n = 0$ formula (25.23) yields

$$0! = \Gamma(1) = 1.$$

Applying (25.22) n times we obtain for $x > 0$

$$\Gamma(x + n) = (x + n - 1)(x + n - 2)\ldots(x + 1)x\Gamma(x). \tag{25.24}$$

(4) $\Gamma(x)$ is *convex down*, i.e., $\Gamma''(x) > 0$. We have

$$\Gamma'(x) = \left(\int_0^{\infty} t^{x-1} e^{-t}\,dt\right)'_x = \int_0^{\infty} t^{x-1} \ln t \, e^{-t}\,dt,$$

and

$$\Gamma''(x) = \int_0^{\infty} t^{x-1} \ln^2 t \cdot e^{-t} dt > 0.$$

It is worth noting that $\Gamma'(x)$ on the ray $(0, +\infty)$ may have only one zero. And since $\Gamma(1) = \Gamma(2) = 1$, by the Rolle theorem this zero x_0 of the derivative $\Gamma'(x)$ exists and lies in the interval $(1, 2)$. Since $\Gamma''(x) > 0$, at x_0 the function $\Gamma(x)$ has a minimum.

It can be shown that $\Gamma(x)$ is differentiable in $(0, +\infty)$ any number of times.

(5) It follows from the formula $\Gamma(x + 1) = x\Gamma(x)$ that

$$\Gamma(x) = \frac{\Gamma(x+1)}{x} \to +\infty$$

as $x \to +0$, because $\Gamma(x)$ is continuous and $\Gamma(x + 1) \to \Gamma(1)$ as $x \to +0$. The plot of the gamma function looks like Fig. 25.4.

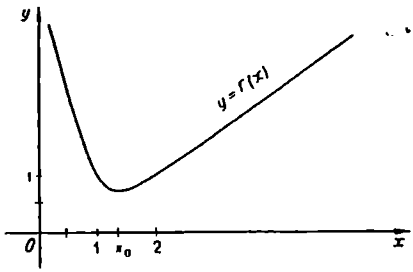


Fig. 25.4

Beta function. The integral

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt \tag{25.25}$$

depending on parameters x and y, is called the *beta function*. The integrand for $x < 1$ and $y < 1$ has two singularities $t = 0$ and $t = 1$.

To find the range of $B(x, y)$ we will represent (25.25) as the sum of two integrals

$$B(x, y) = \int_0^{1/2} t^{x-1}(1-t)^{y-1} dt + \int_{1/2}^{1} t^{x-1}(1-t)^{y-1} dt, \tag{25.26}$$

the first of which for $x < 1$ has the singularity $t = 0$, and the second for $y < 1$ has the singularity $t = 1$.

The integral

$$\int_0^{1/2} t^{x-1}(1-t)^{y-1}\,dt = \int_0^{1/2} \frac{(1-t)^{y-1}}{t^{1-x}}\,dt$$

is an improper integral of the second kind and it converges provided that $1 - x < 1$, i.e., for $x > 0$.

Similarly, the integral

$$\int_{1/2}^{1} t^{x-1}(1-t)^{y-1}\,dt = \int_{1/2}^{1} \frac{t^{x-1}}{(1-t)^{1-y}}\,dt$$

converges for $1 - y < 1$, i.e., for $y > 0$.

So, the beta function $B(x, y)$ is defined for all positive x and y.

We can see that integral (25.25) converges uniformly in each region $x \geqslant a > 0$, $y \geqslant b > 0$, so that the beta function is continuous for $x > 0$, $y > 0$.

Some properties of the beta function. (1) A change $t = z/(z + 1)$ reduces the beta function to the integral

$$B(x, y) = \int_0^{\infty} \frac{z^{x-1}}{(1+z)^{x+y}}\,dz. \tag{25.27}$$

(2) For $x > 0$ and $y > 0$ we have

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \tag{25.28}$$

(3) The beta function is symmetrical in x and y, i.e.,

$$B(x, y) = B(y, x).$$

This follows from (25.28)

(4) A change $t = \sin^2\varphi$, or $t = \cos^2\varphi$, transforms (25.25) to

$$B(x, y) = 2\int_0^{\pi/2} \sin^{2x-1}\varphi\cos^{2y-1}\varphi\,d\varphi$$

$$= 2\int_0^{\pi/2} \cos^{2x-1}\varphi\sin^{2y-1}\varphi\,d\varphi. \tag{25.29}$$

(5) The complement formula is

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}, \qquad 0 < x < 1. \tag{25.30}$$

Euler integrals used to take definite integrals. Euler integrals are adequately examined nonelementary functions. They have been tabulated in much detail. And so it pays to be able to reduce integrals to Euler integrals. We will illustrate this by some examples.

$\int_0^\infty \frac{z^{x-1}}{(1+z)^{x+y}}$

Examples. (1) Compute $\displaystyle\int_0^\infty \frac{\sqrt[4]{t}}{(1+t)^2}\,dt.$

◄ We have

$$\int_0^\infty \frac{\sqrt[4]{t}}{(1+t)^2}\,dt = \int_0^\infty \frac{t^{\frac{5}{4}-1}}{(1+t)^{\frac{5}{4}+\frac{3}{4}}}\,dt = B\left(\frac{5}{4}, \frac{3}{4}\right).$$

We have here used (25.27). If now we apply (25.28) and then (25.23), (25.24) and (25.30), we will obtain

$$B\left(\frac{5}{4}, \frac{3}{4}\right) = \frac{\Gamma\left(\frac{5}{4}\right)\Gamma\left(\frac{3}{4}\right)}{\Gamma(2)} = \frac{1}{4}\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right) = \frac{1}{4}\Gamma\left(\frac{1}{4}\right)\Gamma\left(1-\frac{1}{4}\right)$$
$$= \frac{1}{4}\,\frac{\pi}{\sin \pi/4} = \frac{\pi\sqrt{2}}{4}. \text{ ►}$$

(2) Compute $\displaystyle\int_0^1 \ln^p\left(\frac{1}{x}\right) dx.$

◄ We make a change $1/x = e^t$ or $x = e^{-t}$. Then $dx = -e^{-t}\,dt$, at $x_1 = 0$ we have $t_1 = +\infty$, and at $x_2 = 1$ we get $t_2 = 0$. Therefore

$$\int_0^1 \ln^p\left(\frac{1}{x}\right) dx = \int_0^{+\infty} t^p e^{-t}\,dt = \Gamma(p+1). \text{ ►}$$

(3) Compute $I = \displaystyle\int_0^1 x^{p-1}(1-x^m)^{q-1}\,dx$, where $p, q, m > 0$.

◄ We put $x^m = t$, then $x = t^{1/m}$, $dx = (1/m)t^{(1/m-1)}\,dt$; the limits of integration remain the same so that the integral reduces to the beta function

$$I = \int_0^1 t^{\frac{p-1}{m}}(1-t)^{q-1}\frac{1}{m}\,t^{\frac{1}{m}-1}\,dt$$
$$= \frac{1}{m}\int_0^1 t^{\frac{p}{m}-1}(1-t)^{q-1}\,dt = \frac{1}{m}B\left(\frac{p}{m}, q\right). \text{ ►}$$

(4) Using the equality

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-\frac{1}{2}} e^{-t}\,dt = \sqrt{\pi} \qquad (25.31)$$

take the integral

$$\int_0^1 \sqrt{t - t^2}\,dt.$$

◄ We have

$$\int_0^1 \sqrt{t-t^2}\,dt = \int_0^1 t^{\frac{1}{2}}(1-t)^{\frac{1}{2}}\,dt = \int_0^1 t^{\frac{3}{2}-1}(1-t)^{\frac{3}{2}-1}\,dt$$

$$= B\left(\frac{3}{2}, \frac{3}{2}\right) = \frac{\Gamma^2\left(\frac{3}{2}\right)}{\Gamma(3)} = \frac{\Gamma^2\left(1+\frac{1}{2}\right)}{2}$$

$$= \frac{1}{2}\cdot\frac{1}{4}\Gamma^2\left(\frac{1}{2}\right) = \frac{\pi}{8}.$$

We have here made use of the definition of the beta function and formulas (25.28), (25.22), (25.23) and (25.31). ►

Exercises

Find the limits:

1. $\lim\limits_{y\to 0} \int_0^1 \sqrt[3]{x^4 + y^2}\,dx.$ 2. $\lim\limits_{y\to -\frac{\pi}{2}} \int_{-1}^2 x^3 \cos\left(\frac{\pi}{2} - y\right) dx.$

3. $\lim\limits_{\alpha\to 0} \int_{-1}^1 \sqrt{x^2 + \alpha^2}\,dx.$

Find the derivatives $F'(y)$ of the following functions:

4. $F(y) = \int_{y^2}^{y} \sin(x^2 + y^2)\,dx.$ 5. $F(y) = \int_{y-1}^{y+1} \frac{\sin(xy)}{x}\,dx.$

6. Using the relation

$$\int_0^b \frac{dx}{a^2 + x^2} = \frac{1}{a}\tan^{-1}\frac{b}{a} \quad (a > 0),$$

take the integral

$$\int_0^b \frac{dx}{(a^2 + x^2)^2}.$$

7. Using the relation

$$\int_0^b \frac{dx}{1 + ax} = \frac{1}{a} \ln (1 + ab) \qquad (a > 0)$$

and differentiating with respect to the parameter obtain the following formula:

$$\int_0^b \frac{x\,dx}{(1 + ax)^2} = \frac{1}{a^2} \ln (1 + ab) - \frac{b}{a(1 + ab)}.$$

8. Prove that the integral

$$I(y) = \int_0^\infty \frac{y^2 - x^2}{(x^2 + y^2)^2}\, dx$$

converges uniformly in y on the entire real axis.
9. Prove that the integral

$$I(s) = \int_0^\infty \frac{dx}{x^2 + s^2}$$

converges uniformly in the parameter s on any interval $[\alpha, \beta]$, if $1 \leqslant \alpha < \beta$.
10. Using the relation

$$\int_0^\infty e^{-\alpha x}\, dx = \frac{1}{\alpha} \qquad (\alpha > 0)$$

and differentiating with respect to the parameter take the integral

$$\int_0^\infty x^n e^{-x}\, dx.$$

Using Euler integrals take the following integrals:

11. $\displaystyle\int_0^\infty \frac{dx}{1 + x^3}$. 12. $\displaystyle\int_0^a x^2\sqrt{a^2 - x^2}\, dx$. *Hint*: $\Gamma(1/2) = \sqrt{\pi}$.

13. $\int\limits_0^1 x\sqrt[3]{1-x^3}\,dx.$

Express through Euler integrals:

14. $\int\limits_0^{\pi/2} \sin^m x \cos^n x\,dx.$ 15. $\int\limits_0^{\pi/2} \tan^n x\,dx.$

16. $\int\limits_0^{\infty} x^{2n}e^{-x^2}\,dx$ (n is a positive integer).

17. $\int\limits_0^{\infty} \frac{dx}{(1+x^2)^n}.$

Answers

1. $\frac{3}{7}$. 2. $-\frac{15}{4}$. 3. 1. 4. $F'(y) = 2\int\limits_{y^2}^{y} y\cos(x^2+y^2)\,dx + \sin(2y^2) - 2y\sin(y^2+y^4)$. 5. $F'(y) = \frac{(2y+1)\sin(y^2+y)}{y^2+y} - \frac{(2y-1)\sin(y^2-y)}{y^2-y}$.
6. $\frac{1}{2}\left[\frac{b}{a^2(a^2+b^2)} + \frac{1}{a^3}\tan^{-1}\frac{b}{a}\right]$. 8. Prove that the integral $I(y) = \int\limits_1^{\infty} \frac{y^2-x^2}{(x^2+y^2)^2}\,dx$ converges uniformly on the entire real axis, i.e., for $-\infty < y < +\infty$. First, we notice that $\int \frac{y^2-x^2}{(x^2+y^2)^2}\,dx = \frac{x}{x^2+y^2}$. Second, we make sure that for any $\varepsilon > 0$ the quantity $A(\varepsilon)$ mentioned in the definition of the improper integral that converges uniformly in the parameter y may be $A(\varepsilon) \doteq 1/\varepsilon$. For $B > A$ $\left|\int\limits_B^{\infty} f(x, y)\,dx\right| = \left|\lim\limits_{N\to\infty} \int\limits_B^{N} \frac{y^2-x^2}{(x^2+y^2)^2}\,dx\right| = \left|\lim\limits_{N\to\infty} \frac{x}{x^2+y^2}\Big|_{x=B}^{N}\right| = \left|-\frac{B}{B^2+y^2}\right| \leqslant \frac{1}{B} < \frac{1}{A} = \varepsilon$. But this exactly means that the given improper integral converges uniformly for $-\infty < y < -\infty$.

9. Prove that the integral $I(s) = \int\limits_0^{\infty} \frac{dx}{x^2+s^2}$ converges uniformly for $\alpha \leqslant s \leqslant \beta$, where $\alpha \geqslant 1$. So, for $s \geqslant 1$ $\frac{1}{x^2+s^2} \leqslant \frac{1}{x^2+1}$. Further, $\int\limits_0^{\infty} \frac{dx}{x^2+1} = \lim\limits_{N\to\infty} \tan^{-1} N = \frac{\pi}{2}$, i.e., the integral exists. It follows from this, by the Weierstrass sufficient test, that the integral converges uniformly. 10. We have $\int\limits_0^{\infty} e^{-\alpha x}\,dx = \frac{1}{\alpha}$ $(\alpha > 0)$. Differentiating with respect to the parameter α n times gives $\int\limits_0^{\infty} x^n e^{-\alpha x}\,dx = \frac{n!}{\alpha^{n+1}}$. Hence at $\alpha = 1$

$\int\limits_0^\infty x^n e^{-x}\,dx = n!$ **11.** $\int\limits_0^\infty \frac{dx}{1+x^3} = \left| \begin{matrix} x^3 = t, & x = \sqrt[3]{t}, & dx = \frac{1}{3}t^{-\frac{2}{3}}\,dt \\ x_1 = 0, \ t_1 = 0, & x_2 = \infty, & t_2 = \infty \end{matrix} \right| =$

$\frac{1}{3}\int\limits_0^\infty \frac{t^{-\frac{2}{3}}\,dt}{1+t} = \frac{1}{3}\int\limits_0^\infty \frac{t^{\frac{1}{3}-1}\,dt}{(1+t)^{1/3+2/3}} = \frac{1}{3}B\left(\frac{1}{3}, \frac{2}{3}\right) = \frac{1}{3}\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right) =$

$\frac{1}{3}\Gamma\left(\frac{1}{3}\right)\Gamma\left(1-\frac{1}{3}\right) = \frac{1}{3}\cdot\frac{\pi}{\sin \pi/3} = \frac{2\pi}{3\sqrt{3}}$. ► We have used here formulas (25.27), (25.28) and (25.30). **12.** $\int\limits_0^a x^2\sqrt{a^2-x^2}\,dx = \left| \begin{matrix} x = a\sqrt{t}, \ dx = \frac{a}{2}t^{-\frac{1}{2}}\,dt \\ x_1 = 0, \ t_1 = 0, \ x_2 = a, \ t_2 = 1 \end{matrix} \right| =$

$\int\limits_0^1 a^2 t\sqrt{1-t}\,\frac{a^2}{2}t^{-\frac{1}{2}}\,dt = \frac{a^4}{2}\int\limits_0^1 t^{\frac{1}{2}}(1-t)^{\frac{1}{2}}\,dt = \frac{a^4}{2}B\left(\frac{3}{2}, \frac{3}{2}\right) = \frac{a^4}{2}\,\frac{\Gamma^2\left(\frac{3}{2}\right)}{\Gamma(3)}$

We have used formulas (25.25) and (25.28) to obtain $\Gamma\left(\frac{3}{2}\right) = \Gamma\left(1+\frac{1}{2}\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$, $\Gamma(3) = 2! = 2$. And so $\int\limits_0^a x^2\sqrt{a^2-x^2}\,dx = \frac{\pi a^4}{16}$. **13.** Changing $x^3 = t$, we obtain $\int\limits_0^1 x\sqrt[3]{1-x^3}\,dx = \frac{1}{3}\int\limits_0^1 t^{-\frac{1}{3}}(1-t)^{\frac{1}{3}}\,dt = \frac{1}{3}B\left(\frac{2}{3}, \frac{4}{3}\right) = \frac{1}{3}\,\frac{\Gamma\left(\frac{2}{3}\right)\Gamma\left(\frac{4}{3}\right)}{\Gamma(2)} =$

$\frac{1}{3}\Gamma\left(\frac{2}{3}\right)\Gamma\left(1+\frac{1}{3}\right) = \frac{1}{9}\Gamma\left(\frac{2}{3}\right)\Gamma\left(\frac{1}{3}\right) = \frac{1}{9}\Gamma\left(1-\frac{1}{3}\right)\Gamma\left(\frac{1}{3}\right) = \frac{1}{9}\,\frac{\pi}{\sin \pi/3} =$

$\frac{2\pi}{9\sqrt{3}}$. We have used formulas (25.25), (25.28), and (25.30). **14.** We transform the integral to (25.29) and use formula (25.28) to get $\int\limits_0^{\pi/2} \sin^m x \cos^n x\,dx =$

$$\frac{1}{2}\left[2\int\limits_0^{\pi/2} \sin^{2\frac{m+1}{2}-1} x \cos^{2\frac{n+1}{2}-1} x\,dx\right] = \frac{1}{2}B\left(\frac{m+1}{2}, \frac{n+1}{2}\right) = \frac{1}{2}\,\frac{\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{m+n}{2}+1\right)}.$$

15. Changing $\tan x = \sqrt{t}$ $(t > 0)$ we can readily reduce the integral to the beta function (25.27), namely $\int\limits_0^{\pi/2} \tan^n x\,dx = \int\limits_0^\infty \frac{t^{\frac{n-1}{2}}}{1+t}\,dt = \frac{1}{2}\int\limits_0^\infty \frac{t^{\frac{n+1}{2}-1}}{(1+t)^{\frac{n+1}{2}} + \left(1-\frac{n+1}{2}\right)}\,dt =$

$\frac{1}{2}B\left(\frac{n+1}{2}, 1 - \frac{n+1}{2}\right) = \frac{1}{2}\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(1 - \frac{n+1}{2}\right) = \frac{1}{2}\frac{\pi}{\sin\left(\frac{n+1}{2}\pi\right)} =$

$\frac{1}{2}\frac{\pi}{\sin\left(\frac{\pi}{2} + \frac{n\pi}{2}\right)} = \frac{\pi}{2\cos\frac{n\pi}{2}}$. We have used formulas (25.27), (25.28) and (25.30), and also the relation $\sin(\pi/2 + \alpha) = \cos\alpha$. 16. Putting $x = \sqrt{t}\ (t > 0)$, we reduce the integral to the gamma function (25.19): $\int\limits_0^\infty x^{2n}e^{-x^2}\,dx = \frac{1}{2}\int\limits_0^\infty t^{n-\frac{1}{2}}e^{-t}\,dt =$ $\frac{1}{2}\Gamma\left(n + \frac{1}{2}\right)$. 17. Substituting $x = \sqrt{t}$, we reduce the integral to the beta function (25.27)

$$\int\limits_0^\infty \frac{dx}{(1+x^2)^n} = \frac{1}{2}\int\limits_0^\infty \frac{t^{-\frac{1}{2}}}{(1+t)^n}\,dt = \frac{1}{2}\int\limits_0^\infty \frac{t^{\frac{1}{2}-1}}{(1+t)^{\frac{1}{2}+\left(n-\frac{1}{2}\right)}}\,dt =$$

$$\frac{1}{2}B\left(\frac{1}{2}, n - \frac{1}{2}\right) = \frac{1}{2}\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(n-\frac{1}{2}\right)}{\Gamma(n)} = \frac{\sqrt{\pi}}{2}\frac{\Gamma\left(n-\frac{1}{2}\right)}{(n-1)!}.$$

Chapter 26

Functions of a Complex Variable

26.1 Essentials. Derivative. Cauchy-Riemann Equations

Sets in the complex plane. Let $\varepsilon > 0$ be an arbitrary positive number, and z_0 be an arbitrary complex number.

The set of points z in the complex plane (Fig. 26.1) obeying

$$|z - z_0| < \varepsilon$$

is an open circle of radius ε with centre at z_0.

Now, putting $z_0 = x_0 + iy_0$, $z = x + iy$, we obtain

$$|z - z_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2} < \varepsilon$$

or, squaring,

$$(x - x_0)^2 + (y - y_0)^2 < \varepsilon^2.$$

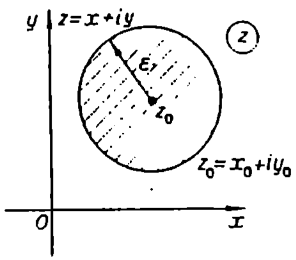


Fig. 26.1

We will refer to the multitude of points z in the complex plane that obey the inequality

$$|z - z_0| < \varepsilon$$

as the *ε-neighbourhood* of z_0.

Definitions. A point z is called an *interior point* of a set in the complex plane, if there exists an ε-neighbourhood of the point such that it wholly belongs to the given set.

A set D is called a *domain* in the complex plain if it has the following properties:

(1) each point in D is an interior point of the set (*openness*);

(2) any two points of the set D can be connected by a broken line consisting of points of this set (*connectedness*).

Any point of the domain D is said to be a *boundary point* if in its any ε-neighbourhood there are points lying both within and without the domain D.

The collection of boundary points, denoted by ∂D, is called the *boundary* of the domain D.

A domain D with the associated boundary ∂D is called a *closed domain* and is denoted by $\bar{D}$.

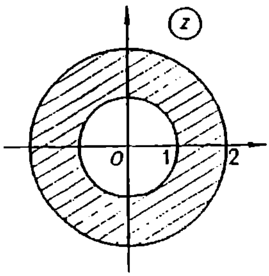


Fig. 26.2

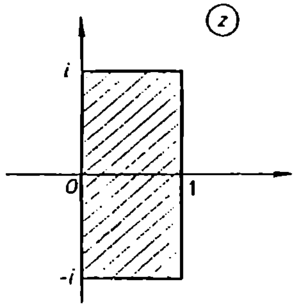


Fig. 26.3

Example. The set of points z (Fig. 26.2) obeying

$$1 < |z| < 2$$

is an *open* domain, and the set obeying

$$1 \leqslant |z| \leqslant 2$$

is a *closed* domain. The boundary consists of two circumferences $|z| = 1$ and $|z| = 2$.

A closed curve without self-intersections will be called a *contour*. Any contour divides the plane into two domains, and is the boundary between them. One of the domains (the *inside* of the contour) is bounded, the other (the *outside*) is unbounded.

We will say that a domain D is *simply connected* if the inside of any contour lying in D also belongs to D.

A domain that is not simply connected will be called *multiply connected*.

Examples. (1) The set of complex numbers $z = x + iy$ subject to the condition

$$0 < x < 1, \qquad -1 < y < 1,$$

is a simply connected domain (Fig. 26.3.)

(2) The set of complex numbers z subject to the condition

$$0 < |z| < 1,$$

is a multiply (doubly) connected domain (Fig. 26.4.) The point $z = 0$ lying inside the contour γ does not belong to the set.

Consider a sequence $\{z_n\}$ of complex numbers

$$z_1, z_2, \ldots, z_n, \ldots .$$

If for an arbitrarily large number $M > 0$ there exists a natural number N such that all the terms z_n of the sequence $\{z_n\}$ with numbers $n > N$ obey the inequality $|z_n| > M$, then we say that the sequence $\{z_n\}$ converges to a *point at infinity*, or simply to infinity, and we write

$$\lim_{n \to \infty} z_n = \infty.$$

If we supplement the plane of the complex variable by the point $z = \infty$ we have thus introduced, we obtain the *extended complex plane*.

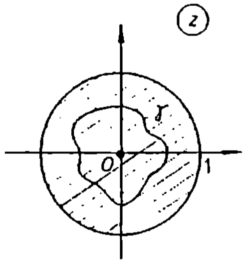


Fig. 26.4

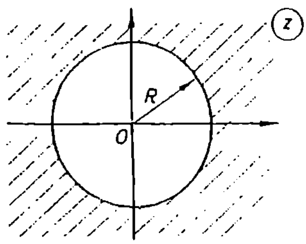


Fig. 26.5

All the points z obeying the inequality $|z| > R$ (plus the point $z = \infty$), i.e., all the points z that lie outside a circle of a sufficiently large radius R with centre at the origin of coordinates, is said to be a *neighbourhood* (R-neighbourhood) of a point at infinity (Fig. 26.5).

Functions of a complex variable. We will say that a function

$$w = f(z)$$

is defined on a set S of the complex plane z if a rule is specified by which each complex number z from S is placed into correspondence with a complex number w (Fig. 26.6).

The function $w = f(z)$ is thus a map of points in the complex plane z onto the complex plane w.

We put $z = x + iy$, $w = u + iv$. Then to define a function of a complex variable $w = f(z)$ will be equivalent to defining two real functions of two

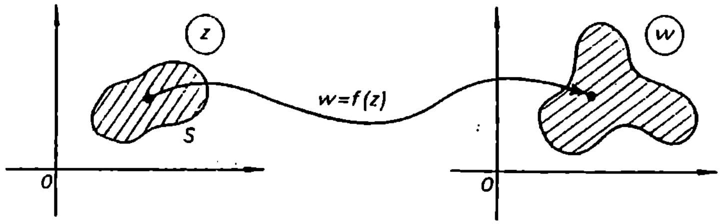


Fig. 26.6

real variables $u = u(x, y)$, $v = v(x, y)$, where

$$w = f(z) = u(x, y) + iv(x, y).$$

Here $u(x, y)$ is said to be the *real part* and $v(x, y)$ the *imaginary part* of the function $w = f(z)$.

Example. Let $w = z^2$.

◄ Putting $z = x + iy$, $w = u + iv$ gives

$$u + iv = (x + iy)^2 = x^2 - y^2 + i\,2xy.$$

And so the expression $w = z^2$ is equivalent to the two expressions $u = x^2 - y^2$, $v = 2xy$. ►

A function $w = f(z)$ is said to be a *univalent* function in a set S, if at different points of this set it assumes different values, otherwise the function is called *multivalent*.

Example. The function $w = z^2$ is univalent in the upper half-plane $\operatorname{Im} z > 0$ and multivalent in the entire plane. For instance, $i^2 = (-i)^2 = -1$.

One often encounters *multi-valued* functions of a complex variable, such that each value of z in S is placed into correspondence with several complex numbers.

Example. The function $w = \sqrt{z}$ is two-valued in the entire plane z, excluding the zero point (and ∞).

Limit of a function. Let a function $w = f(z)$ be defined in a neighbourhood of a point $z_0 = x_0 + iy_0$, except perhaps at z_0 itself.

A complex number A is said to be the *limit* of $f(z)$ when z tends to z_0 (notation: $A = \lim\limits_{z \to z_0} f(z)$), if for any positive ε we can specify a δ-neighbourhood of the point z_0, such that for all points z in it, except perhaps at z_0 itself, the corresponding points w lie in the ε-neighbourhood of the point A (Fig. 26.7).

If z_0 and A are finite points in the complex plane, then we can define the limit as

$$A = \lim_{z \to z_0} f(z), \tag{26.1}$$

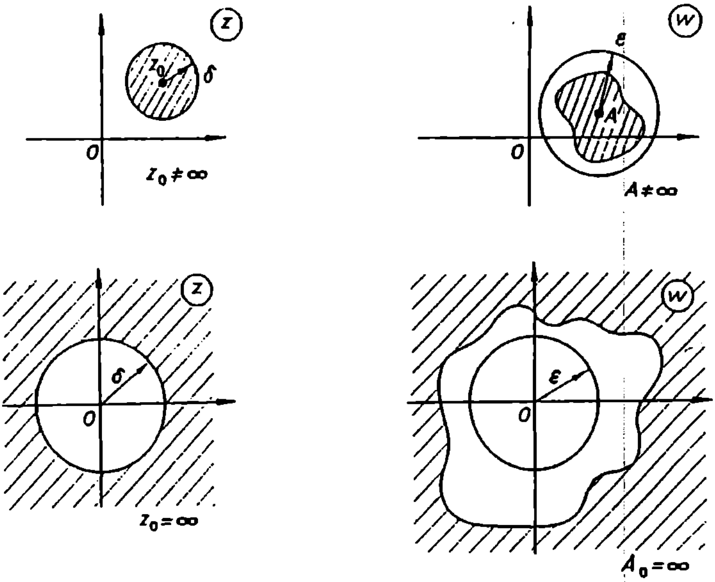


Fig. 26.7

if for any $\varepsilon > 0$, we can specify $\delta = \delta(\varepsilon) > 0$ such that for all z meeting the condition $0 < |z - z_0| < \delta$, we have $|f(z) - A| < \varepsilon$.

It is worth noting that by definition the function $f(z)$ tends to its limit A regardless of the way in which the point z approaches the point z_0.

To say that there exists a limit (26.1) amounts to saying that there exist the limits of the real functions $u(x, y)$ and $v(x, y)$:

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} u(x, y) = B, \qquad \lim_{\substack{x \to x_0 \\ y \to y_0}} v(x, y) = C,$$

where $A = B + iC$.

Because the definition (26.1) reduces to the definition of the limit for real functions of two real variables, the function of a complex variable obeys the following basic limiting relations:

$$\begin{aligned} &\lim_{z \to z_0} f(z) \pm \lim_{z \to z_0} g(z) = \lim_{z \to z_0} (f(z) \pm g(z)), \\ &\lim_{z \to z_0} f(z) \cdot \lim_{z \to z_0} g(z) = \lim_{z \to z_0} f(z) \cdot g(z), \\ &\frac{\lim\limits_{z \to z_0} f(z)}{\lim\limits_{z \to z_0} g(z)} = \lim_{z \to z_0} \frac{f(z)}{g(z)} \qquad (\lim_{z \to z_0} g(z) \neq 0). \end{aligned} \tag{26.2}$$

Continuity. A function $w = f(z)$ defined on a set S is said to be *continuous at a point* $z_0 \in S$, if

$$\lim_{z \to z_0} f(z) = f(z_0), \qquad z \in S.$$

In other words, $f(z)$ is continuous at z_0, if for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all points $z \in S$ meeting the condition $|z - z_0| < \delta$ we have $|f(z) - f(z_0)| < \varepsilon$.

The necessary and sufficient condition for a function of a complex variable $f(x) = u(x, y) + iv(x, y)$ to be continuous at a point $z_0 = x_0 + iy_0$ is that its real and imaginary parts $u(x, y)$ and $v(x, y)$ should be continuous at the point (x_0, y_0) in x and y.

This enables us to translate to functions of a complex variable the basic properties of continuous functions of two real variables, such as the continuity of the sum, product and quotient of two functions, and the continuity of a composite function.

If a function $f(z)$ is continuous at each point of a set S, then we say that $f(z)$ is *continuous on S.*

Differentiability and analyticity. Let a function $f(z)$ be defined in a neighbourhood of a point z. We say that $f(z)$ is *differentiable* at z, if there exists

$$\lim_{h \to 0} \frac{f(z + h) - f(z)}{h}.$$

This limit is called the *derivative* of $f(z)$ at z_0 and is denoted by the symbol $f'(z)$, of $df(z)/dz$:

$$f'(z) = \frac{df(z)}{dz} = \lim_{h \to 0} \frac{f(z + h) - f(z)}{h}. \tag{26.3}$$

It follows from the definition of the derivative (26.3) and the properties of the limit (26.2) that the resulting theory of differentiation of the sum, product and quotient of two functions of a complex variable, of the composite function and inverse function of a complex variable closely resembles the real case:

$$\frac{d}{dz}[f(z) + g(z)] = \frac{df(z)}{dz} + \frac{dg(z)}{dz},$$

$$\frac{d}{dz}[f(z) \cdot g(z)] = \frac{df(z)}{dz} g(z) + f(z) \frac{dg(z)}{dz},$$

$$\frac{d}{dz}\left[\frac{f(z)}{g(z)}\right] = \frac{\frac{df(z)}{dz} g(z) - f(z) \frac{dg(z)}{dz}}{g^2(z)} \qquad (g(z) \neq 0),$$

$$\frac{d}{dz}[f(g(z))] = \frac{df(g(z))}{dw} \frac{dg(z)}{dz}$$

(here $w = g(z)$),

$$\frac{df(z)}{dz} = \frac{1}{d\varphi(w)/dw}, \qquad \frac{d\varphi(w)}{dw} \neq 0$$

(here $z = \varphi(z)$ is a function inverse to $w = f(z)$).

Example. Show that a function $w = f(z) = \operatorname{Re} z$ is not differentiable at any point.

◄ Let $z = z + iy$. Then $w = x$. It will be recalled that by the definition of differentiability of a function $w = f(z)$ at a point z the limit of

$$\frac{f(z + h) - f(z)}{h}$$

must be independent of the way in which one approaches the point z.

Consider two cases.

Let $h = s$ be real. Then

$$\lim_{s \to 0} \frac{f(z + s) - f(z)}{s} = \lim_{s \to 0} \frac{(x + s) - x}{s} = 1.$$

We assume that $h = it$, where t is real. Then

$$\lim_{t \to 0} \frac{f(z + it) - f(z)}{it} = \lim_{t \to 0} \frac{x - x}{it} = 0.$$

It follows that the limiting value of the ratio is influenced by the way in which the point z is approached. Consequently, $w = \operatorname{Re} z$ is not differentiable at any point. ►

The requirement that $f(z)$ be differentiable at a point $z = x + iy$ imposes certain conditions on the behaviour of the real and the imaginary parts of the function in the neighbourhood of the point (x, y).

Theorem 26.1. *Let a function $f(z) = u(x, y) + iv(x, y)$ be differentiable at a point $z = x + iy$. Then at the point (x, y) there exist partial derivatives of $u(x, y)$ and $v(x, y)$ with respect to x and y, such that*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{26.4}$$

The relations (26.4) are called the *Cauchy-Riemann differential equations.*

◄ There exists

$$\lim_{h \to 0} \frac{f(z + h) - f(z)}{h} = f'(z) \tag{26.5}$$

whatever the way in which one approaches the point z.

We suppose first that h tends to zero, while remaining real ($h = s$).

In that case (Fig. 26.8)

$$f'(z) = \lim_{s\to 0} \frac{u(x+s, y) - iv(x+s, y) - u(x, y) - iv(x, y)}{s}$$

$$= \lim_{s\to 0} \frac{u(x+s, y) - u(x, y)}{s} + i \lim_{s\to 0} \frac{v(x+s, y) - v(x, y)}{s}.$$

The last transformation is valid because if a function of a complex variable has a limit it means that there simultaneously exist the limits of its real and imaginary parts.

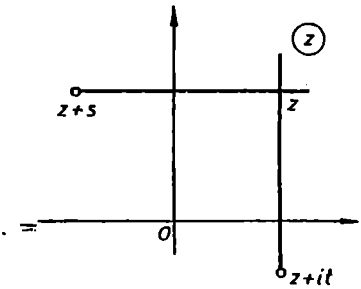


Fig. 26.8

Thereby,

$$f'(z) = \frac{\partial u}{\partial x}(x, y) + i\frac{\partial v}{\partial x}(x, y).$$

If now we put in (26.5) $h = it$, where t is real, we get

$$f'(z) = \lim_{t\to 0} \frac{u(x, y+t) - u(x, y)}{it} + i \lim_{t\to 0} \frac{v(x, y+t) - v(x, y)}{it}$$

$$= -i \lim_{t\to 0} \frac{u(x, y+t) - u(x, y)}{t} + \lim_{t\to 0} \frac{v(x, y+t) - v(x, y)}{t}$$

$$= -i\frac{\partial u}{\partial y}(x, y) + \frac{\partial v}{\partial y}(x, y).$$

In the last two expressions for $f'(z)$ the right-hand sides are equal

$$\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}.$$

And so relation (26.4) is valid. ►

By imposing definite conditions on the real and imaginary parts of a function of a complex variable we can guarantee its differentiability.

Theorem 26.2. *Let the functions $u(x, y)$ and $v(x, y)$ be differentiable at a point (x, y) as functions of real variables. Suppose that at the point*

the conditions (26.4) *are met. Then the function of a complex variable* $f(z) = u(x, y) + iv(x, y)$ *is differentiable at the point* $z = x + iy$.

◀ By definition of differentiability of the real functions $u(x, y)$ and $v(x, y)$ with respect to x and y, their increments at (x, y) can be written as

$$u(x + s, y + t) - u(x, y) = u_x(x, y)s + u_y(x, y)t + \alpha|h|,$$
$$v(x + s, y + t) - v(x, y) = v_x(x, y)s + v_y(x, y)t + \beta|h|$$

(here α and β tend to zero together with $|h| = \sqrt{s^2 + t^2}$).

If we multiply by i the second of these and add together with the first one, we will have

$$f(z + h) - f(z) = (u_x + iv_x)s + (u_y + iv_y)t + \gamma|h|,$$

where $\gamma = \alpha + i\beta$ tends to zero as $h \to 0$. Using (26.4), we will exclude from the above relation u_y and v_y. We can then write the increment of $f(z)$ as follows:

$$f(z + h) - f(z) = (u_x + iv_x)(s + it) + \gamma|h|.$$

Having divided both sides of this by $h = s + it$, we see that

$$\lim_{h \to 0} \frac{f(z + h) - f(z)}{h}$$

does exist and is equal to $u_x + iv_x$. ▶

Examples. (1) The function $w = \bar{z} = x - iy$ is not differentiable at any point, since

$$\frac{\partial u}{\partial x} = 1 \neq \frac{\partial v}{\partial y} = -1. \ ▶$$

A function $w = f(z)$ is said to be *analytic at a point* z, if it is differentiable both at z itself and in some neighbourhood of it.

A function $w = f(z)$ differentiable at each point of some domain D is called an *analytic function in the domain.*

For any analytic function $f(z)$ we have

$$f'(z) = u_x + iv_x = v_y - iu_y = u_x - iu_y = v_y + iv_x. \tag{26.6}$$

(2) Check if the function $w = z\bar{z}$ is analytic at least at one point.

◀ We have $z\bar{z} = x^2 + y^2$, so that

$$u(x, y) = x^2 + y^2, \qquad v(x, y) = 0.$$

The Cauchy-Riemann equations will then be

$$2x = 0, \qquad 2y = 0.$$

They are only met at the point (0, 0).

And so the function $w = z\bar{z}$ is only differentiable at $z = 0$ and is not analytic anywhere. ►

(3) Prove that the function

$$w = f(z) = e^x(\cos y + i \sin y)$$

is analytic in the entire complex plane z.

◄ The functions

$$u(x, y) = e^x \cos y, \qquad v(x, y) = e^x \sin y$$

as functions of real variables x and y are differentiable at any point (x, y). It can readily be checked that their first derivatives meet the conditions (26.4). Using formula (26.6), we will compute the derivative of $f(z)$:

$$f'(z) = (e^x(\cos y + i \sin y))' = e^x(\cos y + i \sin y) = f(z). \text{ ►}$$

Using the Cauchy-Riemann equations, we can restore an analytic function up to a constant, if we know its real part $u(x, y)$ or its imaginary part $v(x, y)$.

(4) Find the analytic function $w = f(z)$ from its real part $u(x, y) = e^x \cos y$ if $f(0) = 1$.

◄ *First method.* Since $u_x = e^x \cos y$, and since $u_x = v_y$, we get $v_y = e^x \cos y$. Hence

$$v(x, y) = \int e^x \cos y \, dy = e^x \sin y + \varphi(x),$$

where $\varphi(x)$ is as yet unknown. Differentiating v with respect to x and using the equality $v_x = -u_y$ gives

$$e^x \sin y + \varphi'(x) = e^x \sin y.$$

It follows that $\varphi'(x) = 0$, and hence $\varphi(x) = C$, where $C = \text{const}$. Since $v(x, y) = e^x \sin y + C$, we have

$$f(z) = e^x \cos y + i(e^x \sin y + C).$$

Further, given $f(0) = 1$, we substitute $x = 0$ and $y = 0$ into the last relation and obtain $1 = 1 + iC$. Hence $C = 0$.

Second method. It is more convenient to seek the imaginary part using a line integral.

We have

$$v_x = -u_y = e^x \sin y, \qquad v_y = u_x = e^x \cos y.$$

Therefore,

$$v(x, y) = \int_{(0,0)}^{(x,y)} e^x \sin y \, dx + e^x \cos y \, dy + C$$

$$= \int_{(0,0)}^{(x,y)} d(e^x \sin y) + C = e^x \sin y + C.$$

Since $v(0, 0) = 0$, we arrive at

$$v(x, y) = e^x \sin y,$$

and

$$f(z) = e^x(\cos y + i \sin y). \blacktriangleright$$

A function $\varphi(x, y)$ is said to be *harmonic* in a domain D, if in the domain it has continuous partial derivatives through the second order and obeys the *Laplace equation*

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} = 0.$$

If a function $f(z) = u + iv$ is analytic in a domain D, then its real part $u(x, y)$ and its imaginary part $v(x, y)$ are harmonic functions in the corresponding region of the xy-plane.

Differentiating the first of (26.4) with respect to x, and the second with respect to y, we will have

$$u_{xx} = v_{yx}, \qquad u_{yy} = -v_{xy}.$$

Using the relation $v_{xy} = v_{yx}$, we arrive at the relation $u_{xx} + u_{yy} = 0$.

A similar relation holds for the imaginary part:

$$v_{xx} + v_{yy} = 0.$$

Remark. The above differentiations have to be justified. Further, in Sec. 26.3, we will prove that a function analytic in a domain has derivatives of all orders in that domain. This of course holds both for its real and its imaginary parts.

Geometrical meaning of the derivative of a function of a complex variable. Let $w = f(z)$ be a function analytic in a domain D. We fix in D a point z_0 and draw through it a smooth curve γ.

Let the function $w = f(z)$ be a map of D in the complex plane z onto a certain domain G of the complex plane w. The point z_0 corresponds to the point w_0 and the curve γ to the curve Γ.

The function $f(z)$ being analytic in D, at each point in D there exists a derivative $f'(z)$. Suppose that $f'(z_0) \neq 0$, and represent the complex number $f'(z_0)$ in exponential form

$$f'(z_0) = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} = \varrho\, e^{i\alpha}.$$

If a point $z = z_0 + \Delta z$ lies on the curve γ, then the corresponding point $w = w_0 + \Delta w$ lies on Γ (Fig. 26.9).

The angle formed by the vector Δz (vector Δw) of the secant curve γ (curve Γ) with the positive x-axis (u-axis) is $\arg \Delta z$ ($\arg \Delta w$). Since in the

limit as $\Delta z \to 0$ and $\Delta w \to 0$ the secants become tangents to respective curves,

$$\arg \Delta z \to \varphi, \qquad \arg \Delta w \to \Phi,$$

where φ (or Φ) is the angle formed by the tangent to γ (or Γ) at z_0 (or w_0) with the x-axis (u-axis).

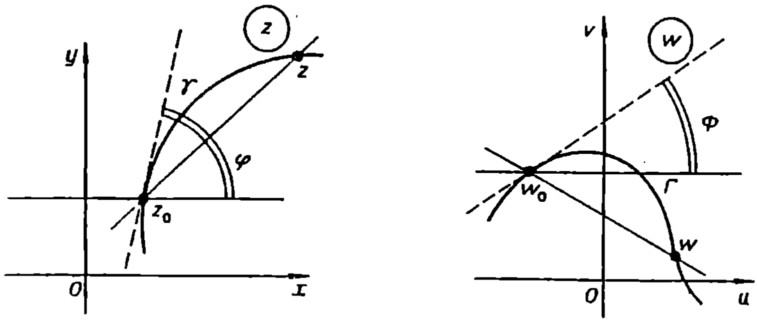


Fig. 26.9

When a complex number is divided by another one, their arguments are subtracted

$$\arg \frac{\Delta w}{\Delta z} = \arg \Delta w - \arg \Delta z.$$

Therefore,

$$\alpha = \arg f'(z_0) = \arg \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z}$$
$$= \lim_{\Delta w \to 0} \arg \Delta w - \lim_{\Delta z \to 0} \arg \Delta z = \Phi - \varphi.$$

Since the value of the derivative is independent of how Δz tends to zero, the resultant difference will be the same for any other smooth curve passing through z_0 (the angles Φ and φ may change, of course).

It follows that in mapping by an analytic function $w = f(z)$ with the derivative $f'(z_0) \neq 0$ the angle

$$\psi = \tilde{\varphi} - \varphi$$

between any two smooth curves γ and $\tilde{\gamma}$ originating from the point z_0 is equal to the angle between their images Γ and $\tilde{\Gamma}$ originating from the point w_0

$$\psi = \tilde{\Phi} - \Phi.$$

Consequently, both magnitudes and directions of the angles remain the same.

This property may be called the *preservation of angles*.

Since

$$\varrho = |f'(z_0)| = \lim_{\Delta z \to 0} \frac{|\Delta w|}{|\Delta z|},$$

we have up to higher-order infinitesimals the relation

$$|\Delta w| = \varrho |\Delta z|,$$

which is independent of the choice of the curve γ.

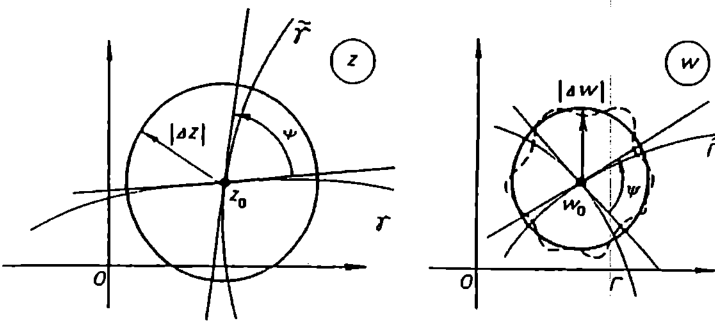


Fig. 26.10

Geometrically this means that infinitesimal circles with centre at z_0 are mapped up to higher-order infinitesimals into small circles with centre at w_0 (Fig. 26.10).

This property is known as that of *constant of extensions*.

A one-to-one mapping $w = f(z)$ of D in the plane z onto the domain G in the plane w is said to be *conformal*, if this mapping at each point of D possesses the two properties just discussed.

The above arguments thus show that a mapping by an analytic function with a nonzero derivative is a conformal one.

Conformity criterion. For a mapping $w = f(z)$ to be conformal in a domain D, it is necessary and sufficient that $f(z)$ be univalent and analytic in D and that $f'(z) \neq 0$ for all z in D.

26.2 Elementary Functions of a Complex Variable

Rational function. A *linear function* of a complex variable z is a function of the form

$$w = az + b, \tag{26.7}$$

where $a \neq 0$ and b are specified complex numbers.

The linear function is defined for all values of the independent variable z. It is single-valued and, since the inverse function

$$z = \frac{1}{a} w - \frac{b}{a} \tag{26.8}$$

is also single-valued, it is univalent in the entire plane z.

The linear function is analytic in the entire complex plane, and, since

$$\frac{dw}{dz} = a \neq 0,$$

the mapping through it is conformal in the entire plane.

A *linear fractional function* is a function of the form

$$w = \frac{az + b}{cz + d}, \tag{26.9}$$

where a, b, c, and d are specified complex numbers, such that

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0.$$

The linear fractional function is defined for all values of the independent variable z, except for $z = -d/c$. It is single-valued and, since the inverse function

$$z = \frac{-dw + b}{cw - a}$$

is single-valued, it is univalent in the entire complex plane, except at the point $z = -d/c$. In this domain function (26.9) is analytic and, since

$$\frac{dw}{dz} = \frac{ad - bc}{(cz + d)^2} \neq 0,$$

the mapping through it is conformal.

We now additionally define (26.9) at $z = -d/c$ by putting $w(-d/c) = \infty$, and make the point $w = \infty$ correspond to the point $z(\infty) = -d/c$. The linear fractional function will then be univalent in the extended complex plane z.

Example. Consider the linear fractional function

$$w = \frac{1}{z}.$$

◄ It follows from the equality $wz = 1$ that the moduli of the complex numbers z and w are related by

$$|w|\,|z| = 1$$

and the numbers themselves lie on the rays originating from the point O symmetrically relative to the real axis.

In particular, the points of the unit circle $|z| = 1$, go into the points of the unit circle $|w| = 1$. This places the complex number

$$z = e^{i\theta}$$

into correspondence with the conjugate number (Fig. 26.11)

$$w = \bar{z} = e^{-i\theta}.$$

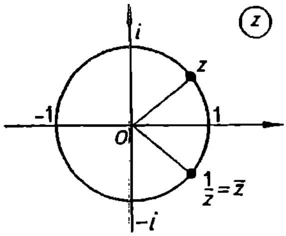


Fig. 26.11

It is to be noted also that the function $w = 1/z$ maps the point $z = \infty$ into $w = 0$. ►

The *power function*

$$w = z^n, \tag{26.10}$$

where n is a natural number, is analytic in the entire complex plane; its derivative $dw/dz = nz^{n-1}$ for $n > 1$ is nonzero at all points save for $z = 0$. If in (26.10) we represent w and z in exponential form

$$w = \varrho e^{i\varphi}, \qquad z = r e^{i\theta},$$

we will find

$$\varrho = r^n, \qquad \varphi = n\theta. \tag{26.11}$$

It is seen from (26.11) that the complex numbers z_1 and z_2 are such that

$$|z_1| = |z_2|, \qquad \arg z_2 = \arg z_1 + \frac{2\pi}{n} k, \tag{26.12}$$

where k is an integer, z_1 and z_2 are going into one point w. This suggests that for $n > 1$ the mapping (26.10) is not univalent in the plane z.

The simplest example of a domain where the mapping $w = z^n$ is univalent is the sector

$$\alpha < \arg z < \alpha + \frac{2\pi}{n}, \tag{26.13}$$

where α is any real number.

In the domain (26.13) the mapping (26.10) is conformal.

Example. The mapping $w = z^n$, $n > 1$, translates the sector $0 < \arg z < \pi/n$ of the plane z into the upper half-plane of w (Fig. 26.12). The mapping increases the sector angle n-fold. Therefore, at the point $z = 0$ the conformity of the mapping $w = z^n$ is violated.

The inverse function is the *nth root*

$$w = \sqrt[n]{z}.$$

It is a multi-valued function, since for each complex number $z = re^{i\theta} \neq 0$ we can specify n different complex numbers

$$w_k = \sqrt[n]{r}\, e^{i\frac{\theta + 2\pi k}{n}} \qquad (k = 0,\ 1,\ \ldots,\ n-1),$$

such that their nth power is z, i.e.,

$$w_k^n = z.$$

Note that $\sqrt[n]{0} = 0$, $\sqrt[n]{\infty} = \infty$.

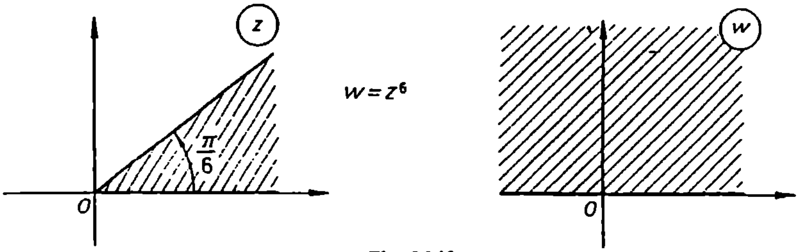


Fig. 26.12

A *polynomial* of degree n of a complex variable z is the function

$$w = a_0 z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n,$$

where $a_0, a_1, \ldots, a_n$ are specified complex numbers ($a_0 \neq 0$). A polynomial of any degree is an analytic function in the entire complex plane.

A *fractional rational* function is a function of the form

$$w = \frac{P(z)}{Q(z)},$$

where $P(z)$ and $Q(z)$ are polynomials of a complex variable z. A fractional rational function is analytic in the entire plane, except for the points where $Q(z) = 0$.

Example. The *Joukowsky function*

$$w = \frac{1}{2}\left(z + \frac{1}{z}\right) \tag{26.14}$$

is analytic in the entire plane z, save for the point $z = 0$.

◄ We now make precise the conditions imposed on the domain in the complex plane, under which the Joukowsky function will be univalent in that domain.

Suppose that (26.14) maps the points z_1 and z_2 into one point. Then

$$\left(z_1 + \frac{1}{z_1}\right) - \left(z_2 + \frac{1}{z_2}\right) = (z_1 - z_2)\left(1 - \frac{1}{z_1 z_2}\right) = 0.$$

At $z_1 \neq z_2$ we find that $z_1 z_2 = 1$.

Therefore, the necessary and sufficient condition for the Joukowsky function to be univalent is

$$z_1 z_2 \neq 1. \tag{26.15}$$

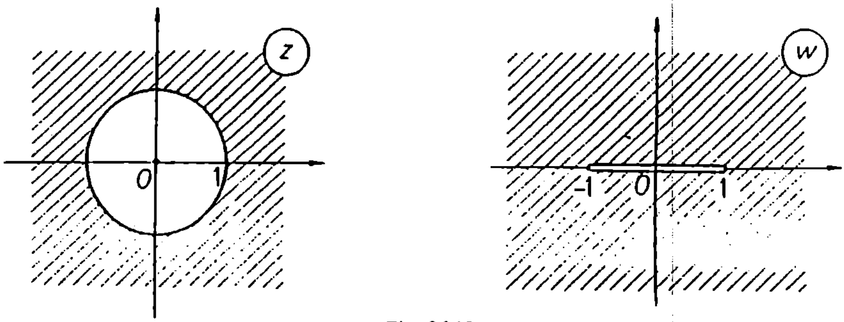


Fig. 26.13

An example of a domain meeting the condition (26.15) is the outside of the circle $|z| > 1$. Since the derivative of the Joukowsky function

$$\frac{dw}{dz} = \frac{1}{2}\left(1 - \frac{1}{z^2}\right)$$

is nonzero everywhere save for the points $z = \pm 1$, then the mapping of the domain $|z| > 1$ by that function will be conformal (Fig. 26.13).

Notice that the inside of the unit circle $|z| < 1$ is also a domain where the Joukowsky function is univalent. ►

We will define the *exponential function* e^z for any complex number $z = x + iy$ by

$$w = e^z = e^{x+iy} = e^x(\cos y + i \sin y). \tag{26.16}$$

At $x = 0$ we obtain the *Euler formula*

$$e^{iy} = \cos y + i \sin y. \tag{26.17}$$

The exponential function has the following *main properties*:

(1) For real z the above definition coincides with the conventional one.

◄ We can verify this directly by putting in (26.16) $y = 0$. ►

(2) The function e^z is analytic in the entire complex plane and it is differentiated using the conventional formula

$$(e^z)' = e^z$$

(see Example (3) on p. 450).

(3) The function satisfies the *law of exponents*

$$e^{z_1} \cdot e^{z_2} = e^{z_1 + z_2}.$$

◄ We put $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$.

$$e^{z_1} \cdot e^{z_2} = e^{x_1}(\cos y_1 + i \sin y_1) e^{x_2} (\cos y_2 + i \sin y_2)$$
$$= e^{x_1 + x_2}(\cos (y_1 + y_2) + i \sin (y_1 + y_2)) = e^{z_1 + z_2}. \text{ ►}$$

(4) The function is periodic with imaginary primitive period $2\pi i$.

◄ For any integral k

$$e^{z + 2\pi k i} = e^z \cdot e^{i2\pi k} = e^z,$$

because

$$e^{i2\pi k} = \cos 2\pi k + i \sin 2\pi k = 1.$$

On the other hand, if $e^{z_1} = e^{z_2}$, where $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, then it follows from the definition (26.16) that

$$e^{x_1} = e^{x_2}, \qquad \cos y_1 = \cos y_2, \qquad \sin y_1 = \sin y_2.$$

And so $x_1 = x_2$, $y_2 = y_1 + 2\pi n$, or

$$z_2 - z_1 = i2\pi n, \tag{26.18}$$

where n is an integer. ►

The band $0 < y < 2\pi$ contains not a single pair of points related by (26.18), therefore the above examination suggests that the mapping $w = e^z$ is univalent in the band $0 < y < 2\pi$. And since the derivative of the function is nonzero ($|e^z| = e^x > 0$), then the mapping is conformal (Fig. 26.14).

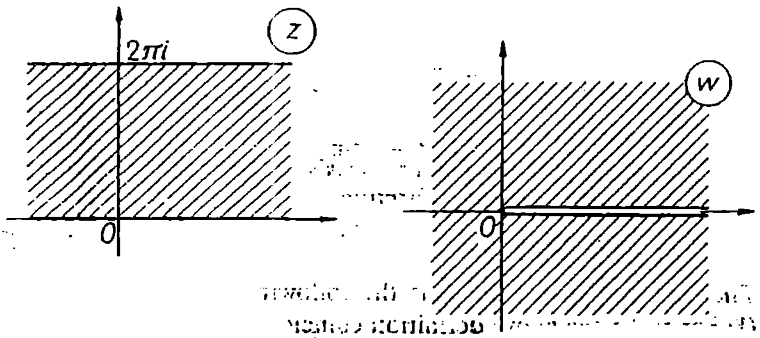


Fig. 26.14

Remark. The function e^z is univalent in any band $\alpha \leqslant y < \alpha + 2\pi$.

Logarithm. From the equation

$$z = e^w,$$

where $z \neq 0$ is specified, and $w = u + iv$ is unknown, we obtain

$$|z| = e^u, \qquad \operatorname{Arg} z = v + 2\pi k \qquad (k = 0, \pm 1, \pm 2, \dots).$$

Hence

$$u = \ln |z|, \qquad v = \operatorname{Arg} z.$$

Thereby the inverse function

$$w = e^z$$

is defined for any $z \neq 0$ and is given by

$$w = \ln |z| + i \operatorname{Arg} z = \ln |z| + i(\arg z + 2\pi k),$$

where $k = 0, \pm 1, \pm 2, \dots$.

This multi-valued function is called the *logarithm* and is denoted by

$$\operatorname{Ln} z = \ln |z| + i \operatorname{Arg} z.$$

The quantity $\ln |z| + i \arg z$ is called the *principal value* of the logarithm. It is given by

$$\ln z = \ln |z| + i \arg z. \tag{26.19}$$

Then

$$\operatorname{Ln} z = \ln z + i2\pi k \qquad (k = 0, \pm 1, \pm 2, \dots). \tag{26.20}$$

Trigonometric and hyperbolic functions. From the Euler formula (26.17) we obtain for real y

$$e^{iy} = \cos y + i \sin y, \qquad e^{-iy} = \cos y - i \sin y.$$

Hence

$$\sin y = \frac{e^{iy} - e^{-iy}}{2i}, \qquad \cos y = \frac{e^{iy} + e^{-iy}}{2}.$$

We define the trigonometric functions $\sin z$ and $\cos z$ for any complex z by

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \qquad \cos z = \frac{e^{iz} + e^{-iz}}{2}. \tag{26.21}$$

The sine and cosine of the complex argument have interesting properties. We will list some of them.

The functions $\sin z$ and $\cos z$:

(1) for real $z = x$ they coincide with the conventional sines and cosines;

(2) are analytic in the entire complex plane;

(3) obey the conventional differentiation formulas:

$$(\sin z)' = \cos z, \qquad (\cos z)' = -\sin z;$$

(4) are periodic with period 2π.

In addition:

(5) $\sin z$ is an odd function, and $\cos z$ is an even function;

(6) they obey the conventional trigonometric relations.

These properties follow readily from (26.21).

The functions $\tan z$ and $\cot z$ in the complex plane are defined by the formulas

$$\tan z = \frac{\sin z}{\cos z}, \qquad \cot z = \frac{\cos z}{\sin z} \tag{26.22}$$

and the *hyperbolic functions* $\sinh z$, $\cosh z$, $\tanh z$, $\coth z$, by the formulas

$$\begin{aligned} \sinh z &= \frac{e^z - e^{-z}}{2}, \qquad \cosh z = \frac{e^z + e^{-z}}{2}, \\ \tanh z &= \frac{\sinh z}{\cosh z}, \qquad \coth z = \frac{\cosh z}{\sinh z}. \end{aligned} \tag{26.23}$$

The hyperbolic functions are closely linked with the trigonometric functions. This connection is given by the following relations:

$$\begin{aligned} \cosh z &= \cos(iz), \qquad \sinh z = -i\sin(iz), \\ \cos z &= \cosh(iz), \qquad \sin z = -i\sinh(iz). \end{aligned} \tag{26.24}$$

The sine and cosine of a complex argument have one more important property: in the complex plane both $|\sin z|$ and $|\cos z|$ assume arbitrarily large positive values. We will show this.

Using the property (6) and formulas (26.24), we obtain

$$\begin{aligned} \sin z &= \sin(x + iy) = \sin x \cos(iy) + \cos(x)\sin(iy) \\ &= \sin x \cosh y - i\cos x \sinh y, \end{aligned}$$

$$\cos z = \cos(x + iy) = \cos x \cosh y + i \sin x \sinh y.$$

Hence

$$\begin{aligned} |\sin z|^2 &= (\sin x \cosh y)^2 + (\cos x \sinh y)^2, \\ |\cos z|^2 &= (\cos x \cosh y)^2 + (\sin x \sinh y)^2. \end{aligned}$$

Putting $x = 0$ gives

$$|\sin z| = |\sinh y|, \qquad |\cos z| = \cosh y.$$

Example. It can be easily seen that

$$\cos(i \ln(5 \pm 2\sqrt{6})) = 5.$$

◄ In fact

$$\cos(i \ln(5 \pm 2\sqrt{6})) = \cosh \ln(5 \pm 2\sqrt{6})$$
$$= \frac{1}{2}(e^{\ln(5+2\sqrt{6})} + e^{\ln(5-2\sqrt{6})})$$
$$= \frac{1}{2}(5 + 2\sqrt{6} + 5 - 2\sqrt{6}) = 5. ►$$

26.3 Integration with Respect to a Complex Argument. Cauchy Theorem. Cauchy Integral Formula

Integral of a function of a complex variable. Given a piecewise smooth oriented curve γ in the complex plane z, we suppose that a function $f(z)$ of a complex variable z is defined on that curve.

We break up the curve γ into n partial arcs

$$z_0 = a, \quad z_1, \ldots, z_{n-1}, z_n = b,$$

where a and b are the ends of the curve γ.

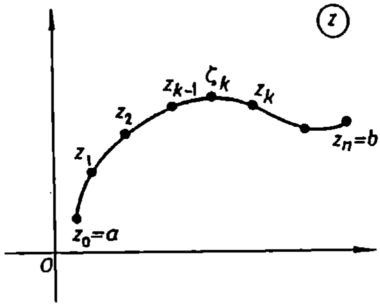


Fig. 26.15

Putting

$$\Delta z_k = z_k - z_{k-1},$$

we form the sum (Fig. 26.15)

$$\sum_{k=1}^{n} f(\zeta_k)\Delta z_k \tag{26.25}$$

(here ζ_k is an arbitrary point of the kth partial arc $[z_{k-1}, z_k]$) called the *complex integral sum* along the curve γ.

If at $\max\limits_k |\Delta z_k| \to 0$ there exists the limit of the sum (26.25) independent of the partition of the curve into partial arcs and of the collection of ζ_k, then this limit is called the *integral* of $f(z)$ along the curve γ:

$$\lim_{\max|\Delta z_k| \to 0} \sum_{k=1}^{n} f(\zeta_k)\Delta z_k = \int_\gamma f(z)\,dz. \tag{26.26}$$

We put

$$f(z) = u(x, y) + iv(x, y),$$

$$z_k = x_k + iy_k, \qquad \Delta x_k = x_k - x_{k-1}, \qquad \Delta y_k = y_k - y_{k-1},$$

$$\zeta_k = \xi_k + i\eta_k, \qquad u_k = u(\xi_k, \eta_k), \qquad v_k = v(\xi_k, \eta_k).$$

We can then write (26.25) as

$$\sum_{k=1}^{n} f(\zeta_k)\,\Delta z_k = \sum_{k=1}^{n} (u_k\Delta x_k - v_k\Delta y_k) + i\sum_{k=1}^{n} (v_k\Delta x_k + u_k\Delta y_k). \tag{26.27}$$

It is seen from this that the real and imaginary parts of (26.25) are integral sums of line integrals of the second kind

$$\int_\gamma u\,dx - v\,dy \qquad \text{and} \qquad \int_\gamma v\,dx + u\,dy, \tag{26.28}$$

respectively.

And so the existence of integral (26.26) depends on the existence of conventional line integrals of functions of real variables. For these integrals to exist it is sufficient for the functions u and v of real variables x and y to be piecewise continuous.

Thus, if γ is a piecewise smooth curve and $f(z)$ is a piecewise continuous function bounded on γ, then integral (26.26) always exists and always holds the formula

$$\int_\gamma f(z)\,dz = \int_\gamma u\,dx - v\,dy + i\int_\gamma v\,dx + u\,dy. \tag{26.29}$$

Formula (26.29) can easily be remembered if written as follows:

$$\int_\gamma f(z)\,dz = \int_\gamma (u + iv)(dx + i\,dy).$$

It follows from (26.29) that integrals of functions of a complex variable retain the basic properties of line integrals of the second kind:

(1) $\int_\gamma f_1(z)\,dz + \int_\gamma f_2(z)\,dz = \int_\gamma [f_1(z) + f_2(z)]\,dz,$

(2) $\int_\gamma cf(z)\,dz = c\int_\gamma f(z)\,dz,$

where c is a complex constant.

(3) $\int\limits_{\gamma^-} f(z)\,dz = -\int\limits_{\gamma^+} f(z)\,dz,$

where γ^- and γ^+ have opposite orientations (Fig. 26.16).

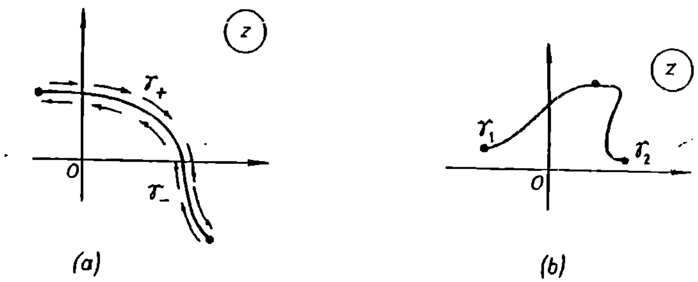


Fig. 26.16

(4) $\int\limits_{\gamma_1} f(z)\,dz + \int\limits_{\gamma_2} f(z)\,dz = \int\limits_{\gamma_1 \cup \gamma_2} f(z)\,dz.$

(5) Let

$$M = \max_{z \in \gamma} |f(z)|$$

and l be the length of the curve γ. Then

$$\left| \int\limits_{\gamma} f(z)\,dz \right| \leqslant \int\limits_{\gamma} |f(z)|\,|dz| \leqslant M \int\limits_{\gamma} |dz| = Ml. \tag{26.30}$$

◄ The proof of (26.30) follows directly from the definition of the integral. When in

$$\left| \sum_{k=1}^{n} f(\zeta_k)\,\Delta z_k \right| \leqslant \sum_{k=1}^{n} |f(\zeta_k)|\,|\Delta z_k| \leqslant M \sum_{k=1}^{n} |\Delta z_k|$$

we pass to the limit as $\max |\Delta z_k| \to 0$ and take into consideration that $\sum\limits_{k=1}^{n} |\Delta z_k|$ is the length of the broken line inscribed into the curve γ, we obtain the required result. ►

Computation of the integral of a function of a complex variable. Let

$$z = z(t) = x(t) + iy(t) \quad (\alpha \leqslant t \leqslant \beta),$$

be a parametric representation of a smooth curve γ. Then we have

$$\int\limits_{\gamma} f(z)\,dz = \int\limits_{\alpha}^{\beta} f[z(t)]z'(t)\,dt. \tag{26.31}$$

◄ Using (26.29) we can reduce the computation of the integral of a function of a complex variable to taking the line integrals (26.28) of real-valued functions. These integrals can be reduced to conventional ones

$$\int_\gamma u\,dx - v\,dy = \int_\alpha^\beta \Big(u(x(t), y(t))x'(t) - v(x(t), y(t)), y'(t)\Big)\,dt,$$
$$\int_\gamma v\,dx + u\,dy = \int_\alpha^\beta \Big(v(x(t), y(t))x'(t) + u(x(t), y(t))y'(t)\Big)\,dt. \tag{26.32}$$

Substituting these into the right-hand side of (26.29), we will obtain the desired result:

$$\begin{aligned}\int_\gamma f(z)\,dz &= \int_\alpha^\beta \Big(u(x(t), y(t)) + iv(x(t), y(t))\Big)(x'(t) + iy'(t))\,dt \\ &= \int_\alpha^\beta f(z(t))\,z'(t)\,dt.\end{aligned}$$ ►

Example. Compute

$$\int_{\gamma_r} \frac{dz}{z - z_0}, \tag{26.33}$$

where γ_r is a circle of radius r with centre at point z_0 traced counterclockwise.

◄ The circle can be represented parametrically as

$$z = z_0 + r\,e^{it}, \qquad 0 \leqslant t < 2\pi.$$

It follows that $z'(t) = ir\,e^{it}$ and

$$\int_{\gamma_r} \frac{dz}{z - z_0} = \int_0^{2\pi} \frac{ir\,e^{it}}{re^{it}}\,dt = i\int_0^{2\pi} dt = 2\pi i.$$

Notice that the value of (26.33) is independent of r and z_0. Reasoning along the same line, we see that

$$I_n = \int_{\gamma_r} (z - z_0)^n\,dz = 0,$$

where n is an integer, $n \neq -1$.

Really,

$$I_n = ir^{n+1}\int_0^{2\pi} e^{i(n+1)t}\,dt = \frac{r^{n+1}}{n+1}\left(e^{2\pi(n+1)i} - 1\right) = 0. ►$$

Theorem 26.3. (Cauchy theorem). *Let a function $f(z)$ be analytic in a simply connected domain D, γ be an arbitrary rectifiable closed curve lying in D. Then*

$$\int_\gamma f(z)\,dz = 0. \tag{26.34}$$

◄ We make two additional assumptions:

(1) γ is a piecewise smooth contour;

(2) $f'(z)$ is continuous.

According to the relation

$$\int_\gamma f(z)\,dz = \int_\gamma u\,dx - v\,dy + i\int_\gamma v\,dx + u\,dy, \tag{26.35}$$

it is sufficient to show that the integrals

$$\int_\gamma u\,dx - v\,dy, \qquad \int_\gamma v\,dx + u\,dy \tag{26.36}$$

are zero.

We denote the inside of γ by G. Since the function $f'(z)$ is continuous everywhere in G, the functions $u(x, y)$ and $v(x, y)$ in that domain have continuous partial derivatives of the first order.

Contour γ being piecewise smooth, we can apply to integrals (26.36) Green's formula,

$$\begin{aligned} \int_\gamma u\,dx - v\,dy &= \iint_G \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy, \\ \int_\gamma v\,dx + u\,dy &= \iint_G \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy. \end{aligned} \tag{26.37}$$

According to the Cauchy-Riemann equations the integrands in each of the double integrals (26.37) are identically equal to zero. ►

From the Cauchy theorem we can make the following remark.

Remark. If a function $f(z)$ is analytic in a simply connected domain D, then the value of the integral

$$\int_\gamma f(z)\,dz,$$

taken along an arbitrary piecewise smooth curve γ lying in D is independent of the choice of the curve γ. It is only determined by the positions of the initial and terminal points of the curve. To stress the independence of the integral $\int_\gamma f(z)\,dz$ of the integration path we will denote it as follows:

$$\int_{z_0}^{z_1} f(z)\,dz,$$

where z_0 and z_1 are the initial and terminal points of γ, respectively.

Theorem 26.4. *Let a function $f(z)$ be analytic in a simply connected domain D, and z_0 and z be points in D. Then the function*

$$F(z) = \int_{z_0}^{z} f(\zeta)\,d\zeta$$

is analytic in D and

$$\frac{dF}{dz} = f(z).$$

◄ By virtue of property (4) of integrals of a function of a complex variable and the previous remark we can represent the relation

$$\frac{F(z+h) - F(z)}{h} = \frac{1}{h}\left(\int_{z_0}^{z+h} f(\zeta)\,d\zeta - \int_{z_0}^{z} f(\zeta)\,d\zeta\right)$$

in the form

$$\frac{F(z+h) - F(z)}{h} = \frac{1}{h}\int_{z}^{z+h} f(\zeta)\,d\zeta. \tag{26.38}$$

We will consider that the integral in (26.38) is taken along a straight segment connecting the points z and $z + h$ (Fig. 26.17).

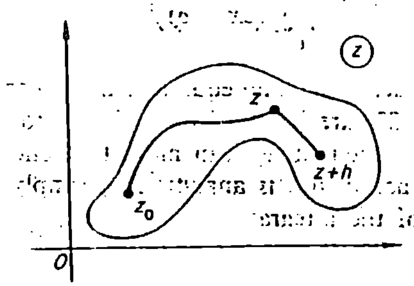


Fig. 26.17

We notice that

$$f(z) = f(z)\frac{1}{h}\int_{z}^{z+h} d\zeta = \frac{1}{h}\int_{z}^{z+h} f(z)\,d\zeta,$$

then

$$\frac{F(z+h)-F(z)}{h} - f(z) = \frac{1}{h}\int_{z}^{z+h} (f(\zeta)-f(z))\,d\zeta.$$

Since $f(\zeta)$ is continuous at z, for any $\varepsilon > 0$ there is $\delta > 0$ such that for $|\zeta - z| < \delta$ will hold the inequality $|f(\zeta) - f(z)| < \varepsilon$.

Let $|h| < \delta$. Then

$$\left|\frac{F(z+h)-F(z)}{h} - f(z)\right| \leqslant \frac{1}{|h|}\int_{z}^{z+h} |f(\zeta)-f(z)|\,|d\zeta|$$
$$< \varepsilon\frac{1}{|h|}\int_{z}^{z+h} |d\zeta| = \varepsilon\frac{|h|}{|h|} = \varepsilon.$$

This suggests that there exists

$$\frac{dF(z)}{dz} = \lim_{h\to 0}\frac{F(z+h)-F(z)}{h} = f(z). \blacktriangleright$$

Remark. It can easily be seen that this treatment is based on two properties of the function $f(z)$:

(1) $f(z)$ is continuous in a domain D;

(2) $\int_{\gamma} f(z)\,dz$ taken along any closed contour γ lying in D is zero, or, which is the same, $\int_{z_0}^{z} f(\zeta)\,d\zeta$ is independent of the integration path.

Under these conditions $F(z) = \int_{z_0}^{z} f(\zeta)\,d\zeta$ is a function analytic in D, such that $F'(z) = f(z)$. We will make use of this note in the next section.

A function $\Phi(z)$ is called the ***antiderivative*** or ***primitive*** of a function $f(z)$ in a domain D, if at each point in D we have

$$\frac{d\Phi(z)}{dz} = f(z). \tag{26.39}$$

We will show that any antiderivative $\Phi(z)$ of $f(z)$ is expressed by

$$\Phi(z) = \int_{z_0}^{z} f(\zeta)\,d\zeta + C, \tag{26.40}$$

where C is a constant, z_0, $z \in D$.

◄ We put

$$w(z) = \Phi(z) - \int_{z_0}^{z} f(\zeta)\,d\zeta = u + iv. \tag{26.41}$$

Then

$$w'(z) = f(z) - f(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y} = 0.$$

It follows that

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} = 0$$

and hence $u(x, y) = C_1$, $v(x, y) = C_2$, where C_1 and C_2 arc constants. Consequently, $w(z) = C_1 + iC_2 = C$.

Putting in (26.41) $z = z_0$, we find that $\Phi(z_0) = C$. ►

Notice that from the relation $\Phi(z_0) = C$ we can write (26.40) in the form

$$\int_{z_0}^{z} f(\zeta)\,d\zeta = \Phi(z) - \Phi(z_0) = \Phi(\zeta)\Big|_{z_0}^{z}. \tag{26.42}$$

And so if $f(z)$ is analytic in a simply connected domain D that contains the points z_0 and z, then as in the real case we have the *Newton-Leibniz formula* (26.42), where $\Phi(z)$ is some antiderivative of $f(z)$.

Exampes. (1) Compute the integral $I = \int_{1-i}^{2+i} (3z^2 + 2z)\,dz$.

◄ The integrand is $f(z) = 3z^2 + 2z$, it is analytic everywhere, its antiderivative is $\Phi(z) = z^3 + z^2$. Using the Newton-Leibniz formula, we find that

$$I = (z^3 + z^2)\Big|_{1-i}^{2+i}$$
$$= (2 + i)^3 + (2 + i)^2 - (1 - i)^3 - (1 - i)^2 = 7 + 19i. \;►$$

(2) Compute the integral $\int_{1}^{z} \frac{d\zeta}{\zeta}$, where $z = re^{i\theta} \neq 0$.

◄ We take as an integration path the piecewise smooth curve $\widehat{ABC}$ (Fig. 26.18) consisting of a segment AB of the real axis with ends at the points 1 and $r = |z|$ and a smaller arc $\widehat{BC}$ of the circle $|\zeta| = r$ with ends

at the points $r = |z|$ and z. Then

$$\int_1^z \frac{d\zeta}{\zeta} = \int_{AB} \frac{d\zeta}{\zeta} + \int_{\widehat{BC}} \frac{d\zeta}{\zeta}.$$

Since $\zeta = x$ in the segment AB, the first integral on the right will be

$$\int_{AB} \frac{d\zeta}{\zeta} = \int_1^r \frac{dx}{x} = \ln r = \ln |z|.$$

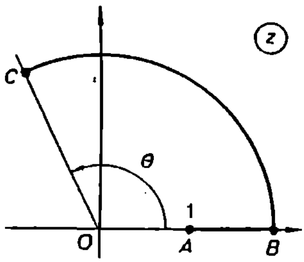


Fig. 26.18

To carry out the second integration on the right we notice that on $\widehat{BC}$ $\zeta = r\,e^{i\varphi}$. Therefore,

$$\int_{\widehat{BC}} \frac{d\zeta}{\zeta} = \int_0^\theta \frac{ir\,e^{i\varphi}\,d\varphi}{r\,e^{i\varphi}} = i\theta = i \arg z.$$

Thus,

$$\int_1^z \frac{d\zeta}{\zeta} = \ln |z| + i \arg z,$$

and hence

$$\int_1^z \frac{d\zeta}{\zeta} = \ln z.$$

On the basis of the theorem just proved we conclude that the principal value of the logarithm $\ln z$ is an analytic function for $z \neq 0$ and that

$$\frac{d \ln z}{dz} = \frac{1}{z}. \blacktriangleright$$

Integration of multi-valued functions. Let a function $w = f(z)$ be analytic in a domain D. It maps D onto a domain G and its inverse function $z = g(w)$ is multi-valued in G. If there exist single-valued (analytic in G) functions $z = g_1(w)$, $z = g_2(w)$, ... for each of which the function $w = f(z)$ is an inverse one, then these functions $g_1(w)$, $g_2(w)$, ... are called *single-valued branches* of $g(w)$ defined in G.

Example. The function $w = z^n$ places into correspondence to each point z a unique point w, but to the same point w ($w \neq 0$, $w \neq \infty$) the function $z = \sqrt[n]{w}$ places into correspondence n various points z (see Chap. 7). If $w = \varrho e^{i\varphi}$, then these n values of z are found by

$$z_k = \sqrt[n]{\varrho}\, e^{i\frac{\varphi + 2\pi k}{n}} \qquad (-\pi < \varphi \leqslant \pi, \qquad k = 0, 1, \ldots, n - 1).$$

Suppose that a simply connected domain G contains a point w_0, but not the points 0 and ∞. Then, with the same choice of a number φ_0 (e.g., $\varphi_0 = \arg w_0$) corresponding to different values of k ($k = 0, 1, \ldots, n - 1$) are different branches of the function $z = \sqrt[n]{w}$. ►

A point is called a *branch point* of a multi-valued function if branches are permuted when one moves round the point in a closed path in a sufficiently small neighbourhood.

For the function $\sqrt[n]{w}$ branch points are $w = 0$ and $w = \infty$. After the point $w = 0$, say, has been traced round n times, we return to the original branch of the function $\sqrt[n]{w}$. Branch points that feature such properties are called *algebraic* branch points (of degree $n - 1$).

At each of these points the function has only one value: $\sqrt[n]{0} = 0$, $\sqrt[n]{\infty} = \infty$, i.e., the various branches of the function coincide at those points.

For $w = \operatorname{Ln} z$, branch points are $z = 0$ and $z = \infty$ ($\operatorname{Ln} 0$ and $\operatorname{Ln} \infty$ being equal to ∞). Any finite number of rounds (in the same direction) about a point $z = 0$ will not lead to the original branch of the function $\operatorname{Ln} z$. Such branch points are called *logarithmic*.

When integrating a multi-valued function one should isolate its single-valued branch. This is achieved by specifying the value of the function at some point on the curve along which the integration is to occur.

Examples. (1) Take the integral $I = \int\limits_\gamma \frac{dz}{\sqrt{z}}$, where γ is the upper semi-circumference $|z| = 1$. For $\sqrt{z}$ take the branch for which $\sqrt{1} = -1$ (Fig. 26.19).

◄ We put $z = re^{i\theta}$, where $r = 1$, and θ varies from 0 to π. From the condition $\sqrt{1} = -1$ it follows that

$$\sqrt{e^{i\theta}} = e^{i\left(\frac{\theta}{2} + \pi\right)}.$$

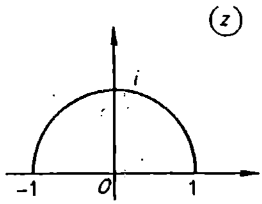


Fig. 26.19

Then

$$I = \int_0^{\pi} \frac{i\,e^{i\theta}}{\sqrt{e^{i\theta}}}\,d\theta = \int_0^{\pi} \frac{i\,e^{i\theta}}{e^{i\left(\frac{\theta}{2}+\pi\right)}}\,d\theta$$

$$= \int_0^{\pi} i\,e^{i\left(\frac{\theta}{2}-\pi\right)}\,d\theta = 2e^{i\left(\frac{\theta}{2}-\pi\right)}\Big|_0^{\pi} = 2\left(e^{-i\frac{\pi}{2}} - e^{-i\pi}\right) = 2(1-i). ►$$

(2) Take the integral $I = \int_1^{i} \frac{\ln^3 z}{z}\,dz$ along the smaller arc of the circle $|z| = 1$ ($\ln z$ is the principal value of the logarithm; $\ln 1 = 0$).

◄ Using the Newton-Leibniz formula, we obtain

$$I = \int_1^{i} \frac{\ln^3 z}{z}\,dz = \int_1^{i} \ln^3 z\,d(\ln z) = \frac{1}{4}\ln^4 z\,\Big|_1^{i}$$

$$= \frac{1}{4}(\ln^4 i - \ln^4 1) = \frac{1}{4}\ln^4 i = \frac{1}{4}\left(\frac{\pi i}{2}\right)^4 = \frac{\pi^4}{64}. ►$$

Cauchy theorem for a multiply connected domain. This Cauchy theorem deals with a contour that wholly lies in a region where the function is analytic. But the theorem also holds for the contour that is the boundary of the region where the function is analytic provided the function is continuous at its closure.

We will now proceed to formulate this generalization of the Cauchy theorem that is important for practical applications.

Theorem 26.5. *Let a function $f(z)$ be analytic in a simply connected domain D and continuous in a closed domain $\bar{D}$, and ∂D be a boundary of the domain. Then*

$$\int_{\partial D} f(z)\,dz = 0. \tag{26.44}$$

In the complex plane we take n closed piecewise smooth contours Γ_0, $\Gamma_1, \ldots, \Gamma_{n-1}$, such that each of the contours $\Gamma_1, \ldots, \Gamma_{n-1}$ lies outside the others and all of them lie inside Γ_0.

The set of points inside Γ_0 and outside $\Gamma_1, \ldots, \Gamma_{n-1}$ is an n-connected domain D (Fig. 26.20).

The complete boundary Γ of D is a composite contour made of the curves $\Gamma_0, \Gamma_1, \ldots, \Gamma_{n-1}$.

We orient the complete boundary Γ of D in the following manner. We say that a multiply connected domain is traced in the positive direction if the domain D is always to the left. The external contour Γ_0 is then traced counterclockwise, and $\Gamma_1, \ldots, \Gamma_{n-1}$ clockwise.

Theorem 26.6. *Let a function $f(z)$ be analytic in a multiply connected domain D and continuous in a closed domain $\bar{D}$. Then*

$$\int_{\Gamma} f(\zeta)\, d\zeta = 0, \tag{26.45}$$

where Γ is the complete boundary of D made up of the contours $\Gamma_0, \Gamma_1, \ldots, \Gamma_{n-1}$ and traced in the positive direction.

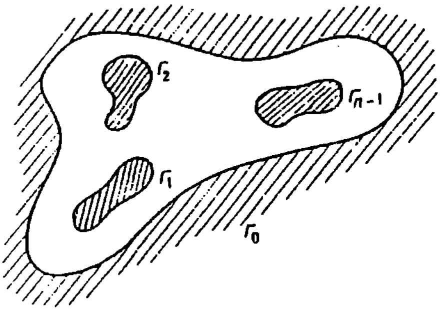


Fig. 26.20

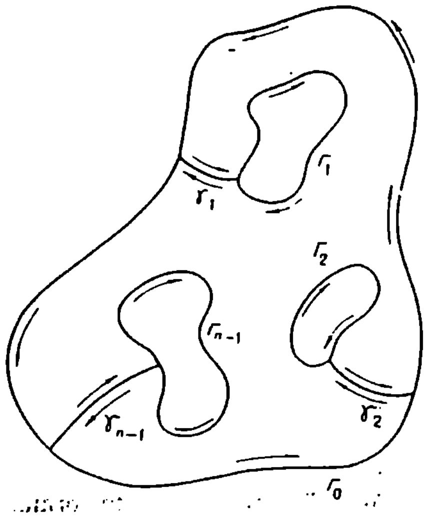


Fig. 26.21

◄ We connect the external contour Γ_0 with $\Gamma_1, \ldots, \Gamma_{n-1}$ by smooth curves $\gamma_1, \ldots, \gamma_{n-1}$, i.e., make cuts, and consider a domain D^* whose boundary Γ^* is made up of the curves $\Gamma_0, \Gamma_1, \ldots, \Gamma_{n-1}$ and curves $\gamma_1, \ldots, \gamma_{n-1}$. The auxiliary curves $\gamma_1, \ldots, \gamma_{n-1}$ are then traced twice in the opposite directions (indicated by arrows in Fig. 26.21); the curves $\gamma_1, \ldots, \gamma_{n-1}$ can always be constructed so that D^* is simply connected.

According to the Cauchy theorem, the integral along the boundary Γ^* of the domain D^* is zero. Since the integrals along γ_k cancel out, then

$$\int_{\Gamma^{*}} f(\zeta)\,d\zeta = \int_{\Gamma_0^{+}} f(\zeta)\,d\zeta + \sum_{k=1}^{n-1} \int_{\Gamma_k^{-}} f(\zeta)\,d\zeta = 0 \tag{26.46}$$

(the superscripts on Γ_k indicate the direction of tracing).

Thus,

$$\int_{\Gamma} f(\zeta)\,d\zeta = 0. ▶$$

Remark. We can rewrite (26.46) as follows

$$\int_{\Gamma_0} f(\zeta)\,d\zeta = \int_{\Gamma_1} f(\zeta)\,d\zeta + \ldots + \int_{\Gamma_{n-1}} f(\zeta)\,d\zeta. \tag{26.47}$$

Theorem 26.7. Cauchy integral formula. *Let a function $f(z)$ be analytic in a domain D and continuous in a closed domain $\bar{D}$. Then, for any internal point z in D we have*

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(\zeta)d\zeta}{\zeta - z}, \tag{26.48}$$

where Γ is the boundary of D traced in the positive direction.

And so *the value of $f(z)$ at any point in D is determined by its values only on the boundary.*

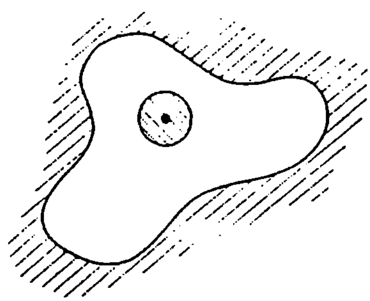


Fig. 26.22

◀ To derive (26.48) we exclude from D a circle of small radius r with centre at a point z (Fig. 26.22). In the domain D^* thus obtained both the numerator and denominator of the integrand $\dfrac{f(\zeta)}{\zeta - z}$ are analytic in ζ, and the denominator is distinct from zero. Therefore, this function is analytic in D^* and continuous in the closed domain $\bar{D}^*$. By the previous theorem the integral along the boundary of D^* is zero, i.e.,

$$\int_{\Gamma} \frac{f(\zeta)}{\zeta - z}\,d\zeta + \int_{\gamma_r^{-}} \frac{f(\zeta)}{\zeta - z}\,d\zeta = 0, \tag{26.49}$$

where γ_r is the circle $|\zeta - z| = r$.

Reversing the direction of integration in the second term gives

$$\int_{\Gamma} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \int_{\gamma_r} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \qquad (26.50)$$

Using the well-known relation

$$\int_{\gamma_r} \frac{d\zeta}{\zeta - z} = 2\pi i$$

(see example on p. 464), we write $f(z)$ as

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(z)}{\zeta - z}\, d\zeta. \qquad (26.51)$$

Dividing both sides of (26.50) by $2\pi i$, we subtract from them $f(z)$. Then we obtain from (26.51) that

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{f(\zeta)}{\zeta - z}\, d\zeta - f(z) = \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta. \qquad (26.52)$$

Notice that the left-hand side of (26.52) is independent of the radius r of the circle excluded. We estimate the right-hand side of the last relation

$$\left| \frac{1}{2\pi i} \int_{\gamma_r} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta \right| \leqslant \frac{1}{2\pi} \int_{\gamma_r} \left| \frac{f(\zeta) - f(z)}{\zeta - z} \right| |d\zeta|$$

$$\leqslant \frac{1}{2\pi} \cdot \frac{\max\limits_{\gamma_r} |f(\zeta) - f(z)|}{r} \cdot 2\pi r = \max_{\gamma_r} |f(\zeta) - f(z)|. \qquad (26.53)$$

Function $f(z)$ is analytic, and hence continuous in D. Therefore, for any $\varepsilon > 0$ we can choose $\delta > 0$ such that $|f(\zeta) - f(z)| < \varepsilon$ for all ζ satisfying the condition $|\zeta - z| < \delta$.

This and (26.53) suggest that by adequately selecting the radius r the integral on the right-hand side of (26.52) can be made as small as we like. On the other hand, the left-hand side of (26.52) is independent of r.

The difference in question is thus zero, which proves the Cauchy formula. ►

If, specifically, Γ is a circle

$$|\zeta - z| = R,$$

then, putting in the Cauchy formula $\zeta - z = R\,e^{i\theta}$, gives

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(z + R\,e^{i\theta})\,d\theta. \tag{26.54}$$

Formula (26.54) is called the *mean value formula.*

Let us formulate the result obtained.

Theorem 26.8. *Let a function $f(z)$ be continuous in a closed circle and analytic inside it. The value of $f(z)$ at the centre of the circle is equal to the mean of its boundary values on the circumference.*

Existence of derivatives of all orders of an analytic function.

Theorem 26.9. *Let a function $f(z)$ be analytic in a domain D and continuous in a closed domain $\bar{D}$. Then at every interior point of D the function has derivatives of all orders*

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_\Gamma \frac{f(\zeta)}{(\zeta - z)^{n+1}}\,d\zeta, \tag{26.55}$$

where Γ is the boundary of D, $n = 1, 2, \ldots$.

◄ To begin with, we make sure that (26.55) holds at $n = 1$. Consider the relation

$$\frac{f(z + h) - f(z)}{h}.$$

Using the Cauchy formula for the values of $f(z)$ at the points z and $z + h$ in D, we write it in the following form

$$\frac{f(z + h) - f(z)}{h} = \frac{1}{h}\,\frac{1}{2\pi i}\left(\int_\Gamma \frac{f(\zeta)}{\zeta - (z + h)}\,d\zeta - \int_\Gamma \frac{f(\zeta)}{\zeta - z}\,d\zeta\right)$$

$$= \frac{1}{2\pi i} \int_\Gamma \frac{f(\zeta)\,d\zeta}{(\zeta - z - h)(\zeta - z)}. \tag{26.56}$$

It can be shown that when $h \to 0$ the function $1/(\zeta - z - h) \to 1/(\zeta - z)$ uniformly for all the points ζ on the curve Γ. Therefore, there exists the limit

$$\lim_{h \to 0} \int_\Gamma \frac{f(\zeta)\,d\zeta}{(\zeta - z - h)(\zeta - z)} = \int_\Gamma \frac{f(\zeta)}{(\zeta - z)^2}\,d\zeta.$$

It follows from this and (26.56) that there exist the derivative of $f(z)$ and

the formula

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h} = \frac{1}{2\pi i} \int_\Gamma \frac{f(\zeta)}{(\zeta - z)^2} d\zeta. \quad (26.57)$$

Assuming (26.55) to be valid for some $k = n$, we can use the same arguments to show that it is valid for $n = k + 1$. ►

Remark. Formula (26.55) can also be proved by differentiating n times the equality

$$f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\zeta)}{\zeta - z} d\zeta \quad (26.58)$$

with respect to z. The right-hand side of (26.58) must be differentiated under the integral sign.

26.4 Complex Power Series. Taylor Series

We recall the basic facts about series (see Chap. 14).

A series of complex numbers

$$c_0 + c_1 + \ldots + c_n + \ldots = \sum_{n=0}^{\infty} c_n, \quad (26.59)$$

where $c_n = a_n + ib_n$ is said to be *convergent* if the sequence of its partial sums

$$\sigma_n = \sum_{k=0}^{n} c_k$$

has a finite limit σ. This limit is called the *sum* of the series (26.59).

Clearly, (26.59) is convergent if and only if the series

$$\sum_{n=0}^{\infty} a_n \quad \text{and} \quad \sum_{n=0}^{\infty} b_n$$

converge simultaneously. Series (26.59) is said to be ***absolutely convergent***, if the series

$$\sum_{n=0}^{\infty} |c_n|$$

is convergent. The series

$$\sum_{n=0}^{\infty} a_n, \quad \sum_{n=0}^{\infty} b_n \quad \text{and} \quad \sum_{n=0}^{\infty} |c_n|$$

are series with real terms and the question of their convergence is solved using the well-known convergence tests for series with real terms.

A *functional series*

$$\sum_{n=0}^{\infty} f_n(z), \tag{26.60}$$

where $f_n(z)$, $n = 0, 1, 2, \ldots$, are defined on some set S of the complex plane, is said to be *convergent* at a point z of the set, if for any $\varepsilon > 0$ there exists N such that for all $n \geqslant N$ we have

$$|R_n(z)| < \varepsilon,$$

where $R_n(z) = \sum_{k=n+1}^{\infty} f_k(z)$.

The series (26.60) is said to be *uniformly convergent* on a set S if (1) it converges at each point of S and (2) for any $\varepsilon > 0$ there is $N = N(\varepsilon)$ independent of z and such that for all $n \geqslant N$ and all z from S the remainders of the series obey

$$\left| \sum_{k=n+1}^{\infty} f_k(z) \right| < \varepsilon.$$

The sufficient test for uniform convergence, which is so important for practical computations, is proved as in the case of one real variable.

Weierstrass test. If everywhere on the set S the series (26.60) is dominated by an absolutely convergent number series

$$|f_n(z)| \leqslant |c_n|,$$

then (26.60) converges on S absolutely and uniformly.

Without any changes we can translate to functions of a complex variable the proof of the proposition that the sum of a uniformly convergent series with continuous terms is continuous; of the theorem stating that a functional series remains uniformly convergent, if its terms are all multiplied by a bounded function; and also of the proposition that a series of continuous functions uniformly convergent on a piecewise smooth curve can be integrated term by term along the curve:

$$\int_{\gamma} \sum_{n=0}^{\infty} f_n(\zeta)\, d\zeta = \sum_{n=0}^{\infty} \int_{\gamma} f_n(\zeta)\, d\zeta. \tag{26.61}$$

Power series. A *power series* is a series of the form

$$c_0 + c_1(z - z_0) + \ldots + c_n(z - z_0)^n + \ldots = \sum_{n=0}^{\infty} c_n(z - z_0)^n, \tag{26.62}$$

where z is an independent complex variable, c_n are specified complex numbers and z_0 is fixed.

Clearly, any power series converges at $z = z_0$.

Examples. (1) The series $\sum_{n=0}^{\infty} n^n z^n$ converges only at $z = 0$.

◄ This follows from the fact that at $z \neq 0$ its nth term $n^n z^n$ does not tend to zero: for any $z \neq 0$, we can find a number, starting with which $|nz| > 2$, and hence $|n^n z^n| \nrightarrow 0$ as $n \to \infty$. ►

(2) The series $\sum_{n=1}^{\infty} \frac{z^n}{n^n}$ converges at each point on the plane.

◄ For each z we can specify a number, starting with which $|z|/n < 1/2$. It follows that $|z|^n/n^n < 1/2^n$, and hence the series is dominated by a convergent number series. ►

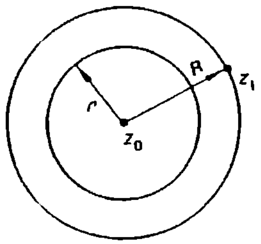


Fig. 26.23

Theorem 26.10 (due to Abel). *Let the power series*

$$\sum_{n=0}^{\infty} c_n(z - z_0)^n \tag{26.63}$$

be convergent at some point $z_1 \neq z_0$ (Fig. 26.23). Then the series

(1) *absolutely converges in the circle*

$$|z - z_0| < |z_1 - z_0| = R,$$

(2) *uniformly converges in the circle*

$$|z - z_0| \leqslant r < R.$$

◄ As stated, the number series

$$\sum_{n=0}^{\infty} c_n(z_1 - z_0)^n \tag{26.64}$$

converges. By the necessary convergence test the nth term of the series (26.64) tends to zero as $n \to \infty$. Any convergent sequence is bounded.

Therefore, there is a constant K such that

$$|c_n(z_1 - z_0)^n| = |c_n|R^n \leqslant K \tag{26.65}$$

for any $n = 0, 1, 2, \ldots$.

We take a point z subject to the condition

$$|z - z_0| \leqslant r < R = |z_1 - z_0| \tag{26.66}$$

and arbitrary otherwise. It follows from this condition that

$$\frac{|z - z_0|}{|z_1 - z_0|} \leqslant \frac{r}{R} = q < 1. \tag{26.67}$$

Thereby, the following estimate holds:

$$|c_n(z - z_0)^n| = |c_n(z_1 - z_0)^n| \cdot \left|\left(\frac{z - z_0}{z_1 - z_0}\right)^n\right| \leqslant Kq^n. \tag{26.68}$$

Inequality (26.68) means that for any point from the circle (26.66) the series (26.62) is dominated by the convergent geometric progression $K\sum_{n=0}^{\infty} q^n$. Consequently, the series (25.62) converges absolutely and uniformly in the circle $|z - z_0| \leqslant r$. ►

Properties of power series. (1) Suppose that power series (26.62) is divergent at some point z_1. Then the series is divergent at each point z that obeys

$$|z - z_0| > |z_1 - z_0|. \tag{26.69}$$

◄ Suppose that the opposite is true: at some point z_2 satisfying the inequality $|z_1 - z_0| < |z_2 - z_0|$ the series converges. Then, by Theorem 26.10 it must converge at z_1 as well. But this is at variance with the conditions. Hence our supposition that there exists a point z_2 with this property is not true. ►

(2) For any power series (26.62) there exists a number R such that in the circle $|z - z_0| < R$ the series (26.62) converges, and outside the circle, for $|z - z_0| > R$, it diverges.

◄ We denote by S a set of points where (26.62) is convergent. The set S is not empty: at $z = z_0$ any series of type (26.62) with any coefficients converges ($R = 0$).

If the set S is unbounded, then (26.62) converges at each point of the complex plane ($R = \infty$).

We suppose that the set S of points where (26.62) converges is bounded. We put

$$R = \sup_{z \in S} |z - z_0|. \tag{26.70}$$

Clearly, at all points z' meeting the inequality $|z' - z_0| > R$ the series (26.62) diverges. ►

If $R > 0$, the largest region of convergence for the series is the circle $|z - z_0| < R$.

On the boundary $|z - z_0| = R$ the series (26.62) may either converge or diverge.

The domain

$$|z - z_0| < R, \qquad R > 0 \tag{26.71}$$

is called the *circle* (or *disc*) *of convergence* of the power series (26.62); the number R given by (26.70) is called the *radius of convergence* of the series.

The convergence radius can be determined from

$$R = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|}, \qquad c_n \neq 0, \tag{26.72}$$

or

$$R = \lim_{n\to\infty} \frac{1}{\sqrt[n]{|c_n|}}, \tag{26.73}$$

if the limits exist (finite or infinite).

Examples. (1) Find the convergence radii of

$$\text{(a)} \quad \sum_{n=0}^{\infty} z^n, \qquad \text{(b)} \quad \sum_{n=1}^{\infty} \frac{z^n}{n^\alpha}, \qquad \alpha > 0.$$

◄ (a) $c_n = 1$. Therefore $\sqrt[n]{|c_n|} = 1$ and $R = 1$.

(b) Here $c_n = 1/n^\alpha$. We now consider

$$\frac{|c_n|}{|c_{n+1}|} = \frac{(n+1)^\alpha}{n^\alpha} = \left(1 + \frac{1}{n}\right)^\alpha \to 1 \qquad \text{as} \qquad n \to \infty.$$

The convergence circle of both series is thus the unit circle $|z| < 1$.

But the sets of points where the series converge are different: series (a) diverges at all points of the circle $|z| = 1$, since the nth term of the series at $|z| = 1$ does not tend to zero; series (b) for $0 < \alpha \leqslant 1$ at certain points of the circle $|z| = 1$ converges (e.g., at $z = -1$), and at others it diverges (e.g., at $z = +1$); for $\alpha > 1$ series (b) converges at all points of the circle absolutely and uniformly, since it is dominated by the convergent number series

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}, \qquad \alpha > 1. ►$$

(2) Prove that the series

$$1 + \frac{z}{1!} + \frac{z^2}{2!} + \ldots + \frac{z^n}{n!} + \ldots = 1 + \sum_{n=1}^{\infty} \frac{z^n}{n!}$$

converges in the entire complex plane z.

◄ According to (26.72),

$$R = \lim_{n \to +\infty} \frac{(n+1)!}{n!} = \lim_{n\to\infty} (n+1) = \infty. ►$$

There is another way of proving this:

◄ Consider the obvious equality

$$(n!)^2 = (1 \cdot n)[2(n-1)]\ldots[(n-1)\cdot 2](n \cdot 1).$$

Each bracket on the right is not smaller than n, since at $k = 1, 2, \ldots$

$$k(n-k+1) - n = (k-1)(n-k) \geqslant 0.$$

Therefore,

$$(n!)^2 > n^n,$$

hence

$$\sqrt[n]{n!} > \sqrt{n}.$$

This implies that

$$\lim_{n\to\infty} \frac{1}{\sqrt[n]{|c_n|}} = \infty \qquad \text{or} \qquad R = \infty. ►$$

In much the same manner, we can prove that the series

$$1 + \sum_{n=1}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} \qquad \text{and} \qquad \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$$

are convergent in the entire complex plane.

Remark. The series

$$\sum_{n=0}^{\infty} z^{n^2}$$

converges in the unit circle $|z| < 1$.

But its convergence radius can be found neither from formula (26.72) ($c_n = 0$, if n is not a square of an integer) nor from formula (26.73) (the terms of the sequence $\sqrt[n]{|c_n|}$ are alternately equal to 1 and 0, and so the sequence has no limit).

In the general case the convergence radius is determined using the *Cauchy-Hadamard formula*

$$R = \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|c_n|}}, \tag{26.74}$$

or

$$\frac{1}{R} = \overline{\lim_{n\to\infty}} \sqrt[n]{|c_n|}.$$

The number l is called the *upper limit* of the sequence of real numbers $\{\alpha_n\}$ and denoted by

$$l = \overline{\lim_{n\to\infty}}\ \alpha_n,$$

if (1) for all $l' > l$ there is a number starting with which all $\alpha_n \leqslant l'$;

(2) there exists a subsequence $\{\alpha_{n_k}\}$ convergent to l.

Each subsequence has an upper limit, finite or infinite.

If the sequence $\{\alpha_n\}$ converges, then $l = \lim\limits_{n\to\infty} \alpha_n$.

Proceeding from the coefficients of the power series

$$c_0 + \sum_{n=1}^{\infty} c_n(z - z_0)^n, \tag{*}$$

construct the sequence of nonnegative numbers

$$|c_1|, \quad \sqrt{|c_2|}, \ \ldots, \quad \sqrt[n]{|c_n|}, \ \ldots .$$

We denote by l the upper limit of the sequence

$$l = \overline{\lim_{n\to\infty}} \sqrt[n]{|c_n|}.$$

Then convergence radius R of the power series (*) is given by the Cauchy-Hadamard formula

$$R = \frac{1}{l}.$$

At $l = 0$ the series (*) is absolutely convergent in the entire plane.

At $l = +\infty$ it converges only at a point z_0 and diverges at $z \neq z_0$.

If $0 < l < +\infty$ the series (*) is absolutely convergent in the circle $|z - z_0| < 1/l$ and diverges in the outside of the circle.

◄ Consider the three cases in succession:

(1) $l = 0$. In this case,

$$\overline{\lim_{n\to\infty}} \sqrt[n]{|c_n|} = \lim_{n\to\infty} \sqrt[n]{|c_n|} = 0.$$

Accordingly, for any z holds the relation

$$\lim_{n\to\infty} \sqrt[n]{|c_n|\,|z - z_0|^n} = 0.$$

By the Cauchy test (see Sec. 13.3) the series

$$\sum_{n=0}^{\infty} |c_n|\,|z - z_0|^n$$

converges, i.e., the series (*) converges absolutely.

(2) $l = +\infty$. There exists a subsequence of numbers $\{n_k\}$ such that

$$\sqrt[n_k]{|c_{n_k}|} \to +\infty.$$

Therefore, for any $z \neq z_0$

$$\sqrt[n_k]{|c_{n_k}|\,|z - z_0|^{n_k}} = \sqrt[n_k]{|c_{n_k}|}\,|z - z_0| \to +\infty.$$

This means that $|c_{n_k}|\,|z - z_0|^{n_k} \to +\infty$. The series (*) thus does not satisfy the necessary convergence test (the nth term of the series does not tend to zero).

(3) $0 < l < +\infty$. If $z = z_0$, all the terms of (*) starting with the second one, vanish, and so the series converges absolutely.

Let $z \neq z_0$ and z lies inside the circle $|z - z_0| < R$. We put $|z - z_0| = \theta^2 R$, where $0 < \theta < 1$. Since $l' = l/\theta > l$, by the definition of the upper limit, beginning with some number all $\sqrt[n]{|c_n|} < l'$. Then

$$\sqrt[n]{|c_n|\,|z - z_0|^n} < l'\,|z - z_0| = \frac{l}{\theta}\,\theta^2 R = \theta < 1.$$

It follows by the Cauchy test that (*) is absolutely convergent.

If z lies outside the circle $|z - z_0| > R$, then $|z - z_0| = R/\theta$, $0 < \theta < 1$. By the definition of the upper limit there exists a sequence of numbers $\{n_k\}$ such that $\sqrt[n_k]{|c_{n_k}|} \to l$.

Therefore,

$$\sqrt[n_k]{|c_{n_k}|\,|z - z_0|^{n_k}} \to l\,|z - z_0| = \frac{lR}{\theta} = \frac{1}{\theta} > 1.$$

Hence, $|c_{n_k}|\,|z - z_0|^{n_k} \to +\infty$, and the series (*) diverges, since it does not satisfy the necessary convergence test. ►

Theorem 26.11 (Taylor theorem). *Let a function $f(z)$ be analytic in a circle $|z - z_0| < R$. Then in the circle $f(z)$ can be represented as a sum of the convergent power series*

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n. \tag{26.75}$$

◄ Let z be an arbitrary point on the circle $|z - z_0| < R$. We construct a circle of radius $r < R$ with centre at point z_0, which contains the point z (Fig. 26.24). We denote by γ_r the circumference $|\zeta - z_0| = r$ that bounds it.

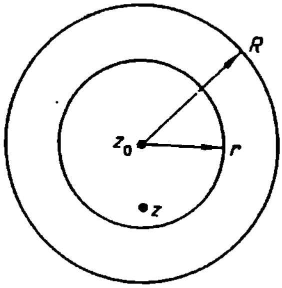


Fig. 26.24

From the Cauchy integral formula we have

$$f(z) = \frac{1}{2\pi i} \int\limits_{\gamma_r} \frac{f(\zeta)}{\zeta - z}\, d\zeta. \tag{26.76}$$

For any point ζ on the circumference γ_r we have

$$\left| \frac{z - z_0}{\zeta - z_0} \right| = \frac{|z - z_0|}{r} = q < 1.$$

The geometric progression

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - z_0)\left(1 - \dfrac{z - z_0}{\zeta - z_0}\right)} = \frac{1}{\zeta - z_0} \sum_{n=0}^{\infty} \frac{(z - z_0)^n}{(\zeta - z_0)^n} \tag{26.77}$$

on γ_r is dominated by the convergent number series

$$\frac{1}{r} \sum_{n=0}^{\infty} q^n = \frac{1}{r(1 - q)}$$

and so it converges absolutely and uniformly in ζ.

We multiply both sides of (26.77) by $\dfrac{1}{2\pi i} f(\zeta)$. This will not violate the uniform convergence of the series, since the function $\dfrac{1}{2\pi i} f(\zeta)$ is continuous, and hence bounded on γ_r. Therefore, it is quite legitimate to inte-

grate term by term the resultant series along γ_r:

$$\frac{1}{2\pi i}\int_{\gamma_r}\frac{f(\zeta)}{\zeta - z}\,d\zeta = \sum_{n=0}^{\infty}\frac{1}{2\pi i}\int_{\gamma_r}\frac{f(\zeta)\,d\zeta}{(\zeta - z_0)^{n+1}}(z - z_0)^n.$$

Putting here

$$c_n = \frac{1}{2\pi i}\int_{\gamma_r}\frac{f(\zeta)\,d\zeta}{(\zeta - z_0)^{n+1}} \quad (n = 0, 1, \ldots) \tag{26.78}$$

and taking into account the Cauchy formula (26.76), we obtain

$$f(z) = \sum_{n=0}^{\infty} c_n(z - z_0)^n. \tag{26.79}$$

Since z is an arbitrary point of the circle $|z - z_0| < R$, it follows from (26.79) that the resultant power series converges to $f(z)$ everywhere inside the circle. ►

Notice that the coefficients c_n are independent of the radius r of the circumference γ_r $(0 < r < R)$.

Power series (26.79), whose coefficients are given by (26.78), is called the *Taylor series* of $f(z)$ with centre at point z_0.

Using formulas for derivatives of analytic functions (see Sec. 32.3), we write the coefficients of the Taylor expansion as

$$c_n = \frac{f^{(n)}(z_0)}{n!} \quad (n = 0, 1, \ldots)$$

and so they are defined uniquely.

Theorem 26.12. *The sum $f(z)$ of the power series*

$$f(z) = \sum_{n=0}^{\infty} c_n(z - z_0)^n \tag{26.80}$$

is analytic in its convergence circle; the derivative $f'(z)$ can be obtained by termwise differentiation

$$f'(z) = \sum_{n=1}^{\infty} nc_n(z - z_0)^{n-1}. \tag{26.81}$$

◄ It is natural to assume that the convergence radius $R > 0$. The power series

$$g(z) = \sum_{n=0}^{\infty} n\,c_n(z - z_0)^{n-1} \text{ and } g(z)(z - z_0) = \sum_{n=1}^{\infty} n\,c_n(z - z_0)^n \tag{26.82}$$

converge and diverge simultaneously. Since

$$\varlimsup_{n\to\infty} \sqrt[n]{n|c_n|} = \varlimsup_{n\to\infty} n^{\frac{1}{n}} \sqrt[n]{|c_n|} = \varlimsup_{n\to\infty} \sqrt[n]{|c_n|},$$

the convergence radius of the series (26.82) is also R.

In each circle U, $|z - z_0| \leqslant r < R$, these series converge uniformly. Accordingly, the function $f(z)$ is continuous in U and the series, whose sum it is, can be integrated term by term. Let γ be an arbitrary contour lying in U. Then

$$\int_\gamma g(z)\,dz = \sum_{n=1}^{\infty} n\,c_n \int_\gamma (z - z_0)^{n-1}\,dz = 0$$

(see Sec. 26.3). By Remark to Theorem 26.4, the function

$$\int_{z_0}^{z} g(\zeta)\,d\zeta = \sum_{n=1}^{\infty} n\,c_n \int_{z_0}^{z} (\zeta - z_0)^{n-1}\,d\zeta = \sum_{n=1}^{\infty} c_n(z - z_0)^n$$

at each point $z \in U$ has a derivative equal to $g(z)$. Then the function

$$f(z) = c_0 + \int_{z_0}^{z} g(\zeta)\,d\zeta$$

will also have the derivative $f'(z) = g(z)$ at each point $z \in U$. ►

Corollary. *Inside the convergence circle a power series can be termwise differentiated and integrated any number of times; the convergence radius of the resultant series will then be equal to the convergence radius of the original series.*

Theorem 26.13. *If a function $f(z)$ can be represented in a circle $|z - z_0| < R$ as a sum of the power series* (26.80), *then the coefficients of the series are determined uniquely by*

$$c_n = \frac{f^{(n)}(z_0)}{n!} \qquad (n = 0, 1, 2, \ldots). \tag{26.83}$$

◄ Putting in (26.80) $z = z_0$ gives

$$f(z_0) = c_0.$$

We now differentiate the series (26.80) termwise

$$f'(z) = c_1 + 2c_2(z - z_0) + \ldots + n\,c_n(z - z_0)^{n-1} + \ldots. \tag{26.84}$$

Setting in (26.84) $z = z_0$, we obtain that

$$f'(z_0) = c_1.$$

We differentiate (26.80) term by term n times

$$f^{(n)}(z) = n!\,c_n + (n + 1)n \ldots 2c_{n+1}(z - z_0) + \ldots.$$

Putting here $z = z_0$ gives

$$f^{(n)}(z_0) = n!\, c_n.$$

To sum up: any convergent power series is a Taylor series of its sum. ►

Formula (26.83) indicates that the coefficients of a Taylor series can be computed as in the real case.

Let us find, for instance, the expansion into a Taylor series with centre at a point $z_0 = 0$ of the functions e^z, $\sin z$ and $\cos z$.

◄ Since

$$\frac{d^n}{dz^n}(e^z)\bigg|_{z=0} = e^z\big|_{z=0} = 1,$$

we have

$$e^z = 1 + \sum_{n=1}^{\infty} \frac{z^n}{n!}.$$

Recall that derivatives of trigonometric functions are found using the same formulas as in the real case. Then

$$\sin z = z - \frac{z^3}{3!} + \ldots + (-1)^n \frac{z^{2n+1}}{(2n+1)!} + \ldots,$$

$$\cos z = 1 - \frac{z^2}{2!} + \ldots + (-1)^n \frac{z^{2n}}{(2n)!} + \ldots.$$

As shown in this section, these series converge in the entire plane.

Example. Find the expansion into a Taylor series with centre at a point $z_0 = 1$ of the function

$$\ln z = \int_1^z \frac{d\zeta}{\zeta}.$$

◄ Using (26.83), we find

$$c_0 = \ln 1 = 0, \qquad c_1 = \frac{1}{z}\bigg|_{z=1} = 1,$$

$$c_n = \frac{1}{n!}(-1)^{n-1}\frac{(n-1)!}{z^n}\bigg|_{z=1} = \frac{(-1)^{n-1}}{n} \qquad (n = 2, 3, \ldots).$$

Hence

$$\ln z = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(z-1)^n}{n}.$$

The resultant series converges in the circle $|z - 1| < 1$ (Fig. 26.25).
These expansions can be found otherwise. Since

$$\frac{1}{1-z} = 1 + z + z^2 + \ldots + z^n + \ldots, \qquad |z| < 1,$$

we have

$$\frac{1}{1+\zeta} = 1 - \zeta + \zeta^2 - \zeta^3 + \ldots + (-1)^n \zeta^n + \ldots.$$

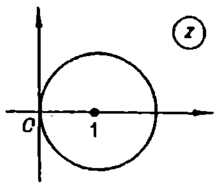


Fig. 26.25

Integrating term by term gives

$$\int_0^z \frac{d\zeta}{1+\zeta} = z - \frac{z^2}{2} + \ldots + \frac{(-1)^n}{(n+1)} z^{n+1} + \ldots$$

or

$$\ln(1 + z) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{z^n}{n}.$$

Putting $1 + z = \zeta$ gives

$$\ln \zeta = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(\zeta - 1)^n}{n}. \quad \blacktriangleright$$

Notice that these expansions can be obtained by formally substituting into the expansions of Chap. 15 the complex variable z for the real variable x.

Cauchy inequalities. Let a function $f(z)$ be analytic in a circle $|z - z_0| < R$ and on the circumference γ_r: $|z - z_0| = r < R$ its modulus being not larger than a constant M.

The coefficients c_n of the Taylor series of $f(z)$ with centre at the point z_0 obey

$$|c_n| \leqslant \frac{M}{r^n} \qquad (n = 0, 1, \ldots). \tag{26.85}$$

◄ As stated

$$|f(\zeta)| \leqslant M$$

for all points ζ on the circumference γ_r. Therefore, using (26.78), we obtain

$$|c_n| \leqslant \frac{1}{2\pi} \int_{\gamma_r} \frac{|f(\zeta)||d\zeta|}{|\zeta - z_0|^{n+1}} \leqslant \frac{1}{2\pi} \frac{M}{r^{n+1}} 2\pi r = \frac{M}{r^n}. \quad ►$$

Theorem 26.14 (Liouville theorem). *Let a function $f(z)$ be analytic in the entire plane, and its modulus be bounded. Then $f(z)$ is constant.*
◄ By the Taylor theorem on any closed circle $|z| \leqslant r$ the function $f(z)$ can be represented in the form of the Taylor series with centre at zero:

$$f(z) = \sum_{n=0}^{\infty} c_n z^n.$$

Since the modulus of $f(z)$ is bounded, i.e.,

$$|f(z)| \leqslant M,$$

the coefficients c_n of the series obey the Cauchy inequalities (26.85). Radius r can be infinitely large. Therefore, for $n = 1, 2, \ldots$ the right-hand sides of (26.85) tend to zero as $z \to \infty$. But the left-hand sides, $|c_n|$, are independent of r. Hence, $c_n = 0$ for $n = 1, 2, \ldots$ and $f(z) = c_0$. ►

Corollary (basic theorem of algebra). *Any polynomial of nonzeroth degree*

$$P(z) = c_0 + c_1 z + \ldots + c_n z^n, \qquad n \geqslant 1, \qquad c_n \neq 0$$

has at least one complex root.
◄ We will prove this by contradiction. Let $P(z)$ have no root. Then

$$f(z) = \frac{1}{P(z)}$$

is an analytic function, subject to the condition

$$\lim_{z \to \infty} f(z) = 0. \tag{26.86}$$

The function $f(x)$ is bounded in magnitude in the entire plane. (From (26.86) follows the existence of $R > 0$ such that for all z, $|z| > R$, holds the inequality $|f(z)| < 1$, if $\max\limits_{|z| \leqslant R} |f(z)| = M$, then $|f(z)| < M + 1$ for all z.) Therefore, by virtue of the Liouville theorem $f(z) = \text{const} = 0$, which is contrary to the definition of $f(z)$. ►

Zeros of analytic function. Let $f(z)$ be an analytic function in a domain D. A point z_0 in D is called a *zero* of $f(z)$, if $f(z_0) = 0$. The expansion

of $f(z)$ about its zero z_0 into a power series has the form

$$f(z) = \sum_{n=1}^{\infty} c_n(z - z_0)^n,$$

i.e., $c_0 = 0$. If the coefficients $c_1, c_2, \ldots, c_{k-1}$ are also equal to zero, and c_k is nonzero, then the point z_0 is called a *zero of order k*. It follows from formula (26.83) that a zero of the kth order is characterized by the relations

$$f(z_0) = 0, \quad f'(z_0) = 0, \ \ldots, \quad f^{(k-1)}(z_0) = 0, \quad f^{(k)}(z_0) \neq 0.$$

In the neighbourhood of a kth-order zero the expansion of $f(z)$ into a power series has the form

$$f(z) = \sum_{n=k}^{\infty} c_n(z - z_0)^n = (z - z_0)^k g(z). \tag{26.87}$$

The function

$$g(z) = \sum_{n=0}^{\infty} c_{n+k}(z - z_0)^n \tag{26.88}$$

has the following properties: it is analytic in the neighbourhood of z_0; $g(z_0) \neq 0$ and the convergence circles of the series (26.87) and (26.88) coincide.

Examples. (1) Find the zeros of $f(z) = 1 - e^z$ and determine their orders.

◄ By equating $f(z)$ to zero we will find the zeros: $z_n = 2\pi ni$, $n = 0, \pm 1, \pm 2, \ldots$. Further,

$$f'(2\pi ni) = -e^{2\pi ni} = -1 \neq 0.$$

Thus, $f(2\pi ni) = 0$, $f'(2\pi ni) \neq 0$. Consequently, the points $z_n = 2\pi ni$ $(n = 0, \pm 1, \pm 2, \ldots)$ are simple zeros of the function $f(z) = 1 - e^z$. ►

(2) Find the order of the zero $z_0 = 0$ of the function

$$f(z) = \frac{z^8}{z - \sin z} \quad (z \neq 0), \qquad f(0) = 0.$$

Using the expansion of $\sin z$ into a Taylor series with centre at the point $z_0 = 0$, we find that

$$f(z) = \frac{z^8}{z - \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} + \ldots\right)} = \frac{z^8}{\frac{z^3}{3!} - \frac{z^5}{5!} + \ldots} = z^5 g(z),$$

where

$$g(z) = \frac{1}{\frac{1}{3!} - \frac{z^2}{5!} + \ldots}.$$

Since $g(0) = 6 \neq 0$, the point $z_0 = 0$ is a zero of the fifth order for this function. ►

26.5 Laurent Series. Isolated Singularities and Their Classification

Taylor series are effective tools of examining functions that are analytic in a circle $|z - z_0| < R$. To examine functions analytic in an annulus

$$0 \leqslant r < |z - z_0| < R \leqslant +\infty$$

it appears possible to construct expansions in positive and negative powers of $(z - z_0)$ of the form

$$\sum_{n=-\infty}^{\infty} c_n(z - z_0)^n \qquad (26.89)$$

that generalize Taylor series.

Series (26.89), understood as a sum of two series

$$\sum_{n=0}^{\infty} c_n(z - z_0)^n, \qquad \sum_{m=1}^{\infty} \frac{c_{-m}}{(z - z_0)^m} \qquad (26.90)$$

is called the *Laurent series.*

Clearly, the convergence domain of series (26.89) is the general part of the convergence domains of each of (26.90). We now proceed to find it.

For the first series

$$\sum_{n=0}^{\infty} c_n(z - z_0)^n \qquad (26.91)$$

the convergence domain is the circle $|z - z_0| < R$, whose radius is determined by the Cauchy-Hadamard formula

$$\frac{1}{R} = \overline{\lim_{n\to\infty}} \sqrt[n]{|c_n|}.$$

Inside the convergence circle series (26.91) converges to an analytic function, and within any circle of a smaller radius $|z - z_0| < R'$, $R' < R$, it converges absolutely and uniformly.

The second series

$$\sum_{m=1}^{\infty} \frac{c_{-m}}{(z - z_0)^m} \tag{26.92}$$

represents a power series relative to the variable $\zeta = 1/(z - z_0)$

$$\sum_{m=1}^{\infty} c_{-m}\zeta^m. \tag{26.93}$$

Series (26.93) converges inside its convergence circle to the analytic function of the complex variable ζ

$$|\zeta| < \frac{1}{r},$$

where

$$r = \overline{\lim_{m\to\infty}} \sqrt[m]{|c_{-m}|}.$$

In any circle of a smaller radius it converges absolutely and uniformly.

This implies that for series (26.92) the domain of convergence is the circle

$$|z - z_0| > r.$$

If $r < R$, there exists a general convergence domain for (26.91) and (26.92), i.e., the *annulus*

$$r < |z - z_0| < R,$$

where series (26.89) converges to an analytic function. In any annulus

$$r' \leqslant |z - z_0| \leqslant R',$$

where $r < r' \leqslant R' < R$, it converges absolutely and uniformly.

Example. Determine the convergence domain for the series

$$\sum_{n=-\infty}^{\infty} c_n z^n = \sum_{n=1}^{\infty} \frac{1}{z^n} + \sum_{n=0}^{\infty} \frac{z^n}{2^{n+1}}.$$

◄ The first series converges in the outside of the circle $|z| > 1$, and the second series converges in the inside of the circle $|z| < 2$. Thereby, the series converges in the annulus $1 < |z| < 2$. ►

Theorem 26.15. *Any function $f(z)$, single-valued and analytic in the annulus $r < |z - z_0| < R$, can be represented in this annulus in the form*

of the sum of the convergent series

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_0)^n \tag{26.94}$$

whose coefficients c_n are defined uniquely and are computed by

$$c_n = \frac{1}{2\pi i} \int_{\gamma_\varrho} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}}\, d\zeta \qquad (n = 0,\ \pm 1,\ \pm 2,\ \ldots), \tag{26.95}$$

where γ_ϱ is the circumference of radius ϱ: $|z - z_0| = \varrho$, $r < \varrho < R$.

◄ We take an arbitrary point z inside the annulus $r < |z - z_0| < R$. We construct the circumferences $\gamma_{r'}$ and $\gamma_{R'}$ with centres at z_0, the radii r' and R' obeying

$$r < r' < |z - z_0| < R' < R.$$

Consider the new annulus $r' < |z - z_0| < R'$ (Fig. 26.26).

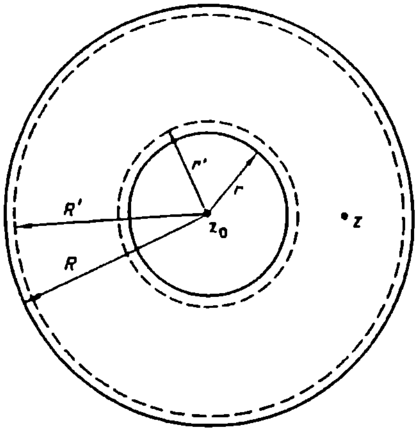


Fig. 26.26

By the Cauchy integral theorem we obtain for the multiply connected domain

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_{R'}^{+}} \frac{f(\zeta)}{\zeta - z}\, d\zeta + \frac{1}{2\pi i} \int_{\gamma_{r'}^{-}} \frac{f(\zeta)}{\zeta - z}\, d\zeta \tag{26.96}$$

We transform individually each of the integrals in the sum (26.96). For all points on the circumference $\gamma_{R'}$ we have

$$\left| \frac{z - z_0}{\zeta - z_0} \right| = q < 1.$$

Therefore, we can represent $1/(\zeta - z)$ as the sum of the uniformly convergent series

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - z_0) - (z - z_0)} = \frac{1}{(\zeta - z_0)\left(1 - \dfrac{z - z_0}{\zeta - z_0}\right)}$$

$$= \frac{1}{\zeta - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{\zeta - z_0}\right)^n.$$

Multiplying both sides by the continuous function $\frac{1}{2\pi i} f(\zeta)$ and integrating term by term along the circumference $\gamma_{R'}$, we obtain

$$\frac{1}{2\pi i} \int_{\gamma_{R'}} \frac{f(\zeta)}{\zeta - z} d\zeta = \sum_{n=0}^{\infty} c_n (z - z_0)^n, \tag{26.97}$$

where

$$c_n = \frac{1}{2\pi i} \int_{\gamma_{R'}} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta. \tag{26.98}$$

We transform the second integral in another way. For all points ζ on $\gamma_{r'}$, we have

$$\left|\frac{\zeta - z_0}{z - z_0}\right| = p < 1.$$

Therefore, we can represent $1/(\zeta - z)$ as the sum of the uniformly convergent series

$$\frac{1}{\zeta - z} = -\frac{1}{z - z_0} \frac{1}{1 - \dfrac{\zeta - z_0}{z - z_0}} = -\frac{1}{z - z_0} \sum_{n=0}^{\infty} \left(\frac{\zeta - z_0}{z - z_0}\right)^n.$$

Multiplying both sides by the continuous function $\frac{1}{2\pi i} f(\zeta)$ and integrating term by term along the circumference $\gamma_{r'}$, we obtain

$$\frac{1}{2\pi i} \int_{\gamma_{r'}} \frac{f(\zeta)}{\zeta - z} d\zeta = \sum_{n=1}^{\infty} \frac{c_{-n}}{(z - z_0)^n}, \tag{26.99}$$

where

$$c_{-n} = -\frac{1}{2\pi i} \int_{\gamma_{r'}} f(\zeta)(\zeta - z_0)^{n-1} d\zeta. \tag{26.100}$$

Notice that the integrands in (26.98) and (26.100) are analytic functions in the annulus $r < |z - z_0| < R$. Therefore, by the Cauchy theorem the values of the integrals will not change, if we replace the circles $\gamma_{R'}$ and $\gamma_{r'}$ by any circle γ_ϱ: $|z - z_0| = \varrho$, where $r < \varrho < R$. This enables us to combine formulas (26.98) and (26.100):

$$c_n = \frac{1}{2\pi i} \int\limits_{\gamma_\varrho} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta \qquad (n = 0, \pm 1, \pm 2 \ldots). \quad (26.101)$$

Substituting (26.97) and (26.99) for the respective integrals on the right of (26.96), we will obtain the expansion required

$$f(z) = \sum_{n=0}^{\infty} c_n(z - z_0)^n + \sum_{n=1}^{\infty} \frac{c_{-n}}{(z - z_0)^n} = \sum_{n=-\infty}^{\infty} c_n(z - z_0)^n. \quad (26.102)$$

Since z is an arbitrary point of the circle $r < |z - z_0| < R$, it follows that series (26.102) converges to the function $f(z)$ everywhere in the annulus, and in any annulus $r < r' \leqslant |z - z_0| \leqslant R' < R$ the series converges to this function absolutely and uniformly.

We prove now that the expansion of the type (26.94) is unique. Suppose that we have another expansion

$$f(z) = \sum_{n=-\infty}^{\infty} c_n'(z - z_0)^n.$$

Then inside the annulus $r < |z - z_0| < R$ we will have

$$\sum_{n=-\infty}^{\infty} c_n'(z - z_0)^n = \sum_{n=-\infty}^{\infty} c_n(z - z_0)^n. \quad (26.103)$$

On the circumference γ_ϱ: $|z - z_0| = \varrho$, $r < \varrho < R$, series (26.103) converges uniformly. We multiply both sides of (26.103) by $(z - z_0)^{-m+1}/2\pi i$, where m is a fixed integer, and integrate both series term by term.

From (26.55) we obtain c_m' on the left, and c_m on the right. Thus, $c_m' = c_m$. Since m is an arbitrary number, then the last equality proves the uniqueness of the expansion. ►

Series (26.94), whose coefficients are found by (26.95), is called the *Laurent series* of $f(z)$ in the annulus $r < |z - z_0| < R$.

The collection of the terms of the series with nonnegative powers of $(z - z_0)$ is called the *regular part* of the Laurent series, and with negative ones, its *principal part*.

Formulas (26.95) for the coefficients of the Laurent series are used only rarely, because as a rule they require cumbersome calculations. Ge-

nerally, if possible, ready-made Taylor expansions of elementary functions are employed. Since the expansion is unique, any legitimate procedure leads to the same result.

Examples. (1) Consider various Laurent expansions of the function

$$f(z) = \frac{2z+1}{z^2+z-2}$$

at $z_0 = 0$.

◄ The function $f(z)$ has two singularities: $z_1 = -2$ and $z_2 = 1$. Consequently, there are three annuli with centre at $z_0 = 0$, where $f(z)$ is analytic:

(a) the annulus $|z| < 1$, (b) the annulus $1 < |z| < 2$ and (c) the outside of the annulus $|z| > 2$ (Fig. 26.27).

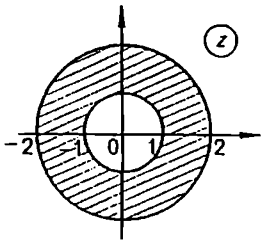


Fig. 26.27

We find the Laurent expansions of $f(z)$ in each of the domains.

We represent $f(z)$ as the sum of elementary fractions

$$f(z) = \frac{1}{z+2} + \frac{1}{z-1} \tag{26.104}$$

(a) the circle $|z| < 1$. We transform relation (26.104) in the following manner:

$$f(z) = \frac{1}{z+2} + \frac{1}{z-1} = \frac{1}{2}\frac{1}{1+\frac{z}{2}} - \frac{1}{1-z}. \tag{26.105}$$

Using the formula for the sum of a geometric progression (Chap. 14), we obtain

$$\frac{1}{1+\frac{z}{2}} = 1 - \frac{z}{2} + \frac{z^2}{4} - \frac{z^3}{8} + \ldots, \qquad |z| < 2, \tag{26.106}$$

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \ldots, \qquad |z| < 1. \tag{26.107}$$

Substituting these into (26.105) gives

$$\frac{2z+1}{z^2+z-2} = \frac{1}{2} - \frac{z}{4} + \frac{z^2}{8} - \frac{z^3}{16} + \dots - (1 + z + z^2 + z^3 + \dots)$$
$$= -\frac{1}{2} - \frac{5}{4}z - \frac{7}{8}z^2 - \frac{17}{16}z^3 - \dots, \qquad |z| < 1.$$

This expansion is the Taylor series of $f(z)$.

(b) the annulus $1 < |z| < 2$. Series (26.106) for the function $1/(1 + z/2)$ remains convergent in the annulus, since $|z| < 2$. Series (26.107) for the function $1/(1 - z)$ for $|z| > 1$ diverges. Therefore, we will transform $f(z)$ as follows:

$$f(z) = \frac{1}{2}\frac{1}{1+\frac{z}{2}} + \frac{1}{z}\frac{1}{1-\frac{1}{z}}. \tag{26.108}$$

Using (26.107) again gives

$$\frac{1}{1-\frac{1}{z}} = 1 + \frac{1}{z} + \frac{1}{z^2} + \frac{1}{z^3} + \dots. \tag{26.109}$$

This series converges for $|1/z| < 1$, i.e., for $|z| > 1$. Substituting expansions (26.106) and (26.109) into (26.108) gives

$$\frac{2z+1}{z^2+z-2} = \frac{1}{2} - \frac{z}{4} + \frac{z^2}{8} - \frac{z^3}{16} + \dots + \frac{1}{z} + \frac{1}{z^2} + \dots$$
$$= \dots + \frac{1}{z^2} + \frac{1}{z} + \frac{1}{2} - \frac{z}{4} + \frac{z^2}{8} - \frac{z^3}{16} + \dots,$$

or

$$\frac{2z+1}{z^2+z-2} = \sum_{n=1}^{\infty}\frac{1}{z^n} + \sum_{n=0}^{\infty}\frac{z^n}{2^{n+1}}, \qquad 1 < |z| < 2.$$

(b) the outside of the circle $|z| > 2$. Series (26.106) for $1/(1 + z/2)$ when $|z| > 2$ diverges, and series (26.109) for $1/(1 - 1/z)$ converges.

We represent $f(z)$ as follows:

$$f(z) = \frac{1}{z}\frac{1}{1+\frac{2}{z}} + \frac{1}{z}\frac{1}{1-\frac{1}{z}} = \frac{1}{z}\left(\frac{1}{1+\frac{2}{z}} + \frac{1}{1-\frac{1}{z}}\right).$$

Using (26.106) and (26.107) gives

$$f(z) = \frac{1}{z}\left(1 - \frac{2}{z} + \frac{4}{z^2} - \frac{8}{z^3} + \dots + 1 + \frac{1}{z} + \frac{1}{z^2} + \frac{1}{z^3} + \dots\right)$$
$$= \frac{1}{z}\left(2 - \frac{1}{z} + \frac{5}{z^2} - \frac{7}{z^3} + \dots\right),$$

or

$$\frac{2z + 1}{z^2 + z - 2} = \frac{2}{z} - \frac{1}{z^2} + \frac{5}{z^3} - \frac{7}{z^4} + \ldots, \qquad |z| > 2. \;\blacktriangleright$$

This example illustrates that the Laurent expansion of a function $f(z)$ will, generally speaking, be different for different annuli.

(2) Find the Laurent expansion of the function $f(z) = \dfrac{2z + 1}{z^2 + z - 2}$ in the annulus $0 < |z - 1| < 3$.

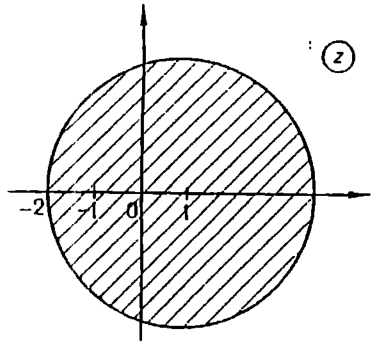


Fig. 26.28

◄ We represent $f(z)$ as follows (Fig. 26.28):

$$f(z) = \frac{1}{z - 1} + \frac{1}{z + 2} \tag{26.110}$$

and transform the second term

$$\frac{1}{z + 2} = \frac{1}{3 + (z - 1)} = \frac{1}{3} \frac{1}{1 - \dfrac{z - 1}{3}}.$$

Using the formula for the sum of a geometric progression, we will obtain

$$\frac{1}{1 + \dfrac{z - 1}{3}} = 1 - \frac{z - 1}{3} + \frac{(z - 1)^2}{9} - \frac{(z - 1)^3}{27} + \ldots,$$

$$|z - 1| < 3.$$

Substituting the expressions obtained into (26.110) gives

$$f(z) = \frac{1}{z - 1} + \frac{1}{3} - \frac{z - 1}{9} + \frac{(z - 1)^2}{27} - \frac{(z - 1)^3}{81} + \ldots \;\blacktriangleright$$

(3) Find the Laurent expansion of the function $f(z) = z^2 \cos \frac{1}{z}$ in a neighbourhood of the point $z_0 = 0$.

◄ For any complex ζ, we have

$$\cos \zeta = 1 - \frac{\zeta^2}{2!} + \frac{\zeta^4}{4!} - \frac{\zeta^6}{6!} + \ldots.$$

We put $\zeta = 1/z$. Then

$$z^2 \cos \frac{1}{z} = z^2 \left(1 - \frac{1}{2!z^2} + \frac{1}{4!z^4} - \frac{1}{6!z^6} + \ldots\right)$$

$$= z^2 - \frac{1}{2!} + \frac{1}{4!z^2} - \frac{1}{6!z^4} + \ldots.$$

This expansion is valid for any point $z \neq 0$. In the case, the annulus is the entire complex plane with one point $z = 0$ punctured out. We can define the region by

$$0 < |z - 0| < \infty.$$

Here $r = 0$, $R = \infty$, $z_0 = 0$.

This function is analytic in the domain $|z| > 0$. ►

Arguing along the same lines as in the previous section, we can obtain from (26.101) for the coefficients of the Laurent series the *Cauchy inequalities*: if a function $f(z)$ is bounded on the circumference $|z - z_0| = \varrho$ ($|f(z)| \leqslant M$, where M is constant), then

$$|c_n| \leqslant \frac{M}{\varrho^n} \qquad (n = 0, \pm 1, \pm 2, \ldots).$$

Isolated singularities. A point z_0 is called an *isolated singularity* of a function $f(z)$, if there exists an annulus-type neighbourhood of z_0

$$0 < |z - z_0| < \varepsilon^{*)}$$

on which $f(z)$ is a single-valued and analytic.

Depending on the behaviour of $f(z)$, as one approaches z_0, one may encounter three types of singularities.

An isolated singularity may be:

(1) *removable*, if there exists a finite $\lim\limits_{z \to z_0} f(z)$;

(2) *a pole*, if $\lim\limits_{z \to z_0} f(z) = \infty$;

(3) *essential*, if $f(z)$ has no limit as $z \to z_0$.

---

*) This set is called the *punctured neighbourhood* of z_0.

The type of an isolated singularity depends on the nature of the Laurent expansion of $f(z)$ in the circle $0 < |z - z_0| < \varepsilon$ with the punctured centre z_0.

Theorem 26.16. *An isolated singularity z_0 of a function $f(z)$ is a removable singularity if and only if the Laurent expansion of $f(z)$ in a neighbourhood of z_0 contains no principal part, i.e., has the form*

$$f(z) = \sum_{n=0}^{\infty} c_n(z - z_0)^n. \qquad (26.111)$$

◄ Let z_0 be a removable singularity. There exists a finite $\lim_{z \to z_0} f(z)$, and hence $f(z)$ is bounded in the punctured neighbourhood of the point z_0: $0 < |z - z_0| < \varepsilon$. We put $|f(z)| \leqslant M$. By the Cauchy inequality

$$|c_n| \leqslant M\varrho^{-n} \qquad (n = 0, \pm 1, \pm 2, \ldots).$$

Since ϱ can be as small as you like, all the coefficients at negative powers of $(z - z_0)$ are equal to zero: $c_n = 0$, $n = -1, -2, \ldots$.

Consequently, if the Laurent expansion of $f(z)$ about z_0 only contains a regular part, i.e., it has the form (26.111), and hence is a Taylor expansion, we can then easily see that as $z \to z_0$ there exists the limiting value of $f(z)$

$$\lim_{z \to z_0} f(z) = c_0. \text{ ►}$$

Theorem 26.17. *An isolated singularity z_0 of a function $f(z)$ is removable if and only if $f(z)$ is bounded in some punctured neighbourhood of z_0: $0 < |z - z_0| < \varepsilon$.*

◄ Let z_0 be a removable singularity of $f(z)$. Putting

$$f(z_0) = \lim_{z \to z_0} f(z)$$

we will find that $f(z)$ is analytic in some circle with centre at z_0. ►

This explains the name of the singularity—*removable*.

Theorem 26.18. *An isolated singularity z_0 of a function $f(z)$ is a pole if and only if the principal part of the Laurent expansion of $f(z)$ in a neighbourhood of the point contains a finite (and positive) number of nonzero terms, i.e., it has the form*

$$f(z) = \sum_{n=-m}^{\infty} c_n(z - z_0)^n, \qquad m > 0, \qquad c_{-m} \neq 0. \qquad (26.112)$$

◄ Let z_0 be a pole. Since $\lim_{z \to z_0} f(z) = \infty$, there exists a punctured neighbourhood of z_0, where $f(z)$ is analytic and nonzero. Then in the neigh-

bourhood the analytic function

$$g(z) = \frac{1}{f(z)}$$

is defined such that

$$\lim_{z \to z_0} g(z) = \frac{1}{\lim\limits_{z \to z_0} f(z)} = 0.$$

Accordingly, the point z_0 is a removable singularity (zero) of $g(z)$ and

$$g(z) = b_m(z - z_0)^m + b_{m+1}(z - z_0)^{m+1} + \ldots, \quad b_m \neq 0, \; m \geqslant 1,$$

or

$$g(z) = (z - z_0)^m h(z),$$

where $h(z)$ is an analytic function; $h(z_0) \neq 0$. Then

$$f(z) = \frac{1}{g(z)} = \frac{1}{(z - z_0)^m} \frac{1}{h(z)}.$$

By the relation $h(z_0) \neq 0$, the function $1/h(z)$ is analytic in a neighbourhood of the point z_0, and hence

$$\frac{1}{h(z)} = \sum_{n=0}^{\infty} c_{n-m}(z - z_0)^n.$$

It follows that

$$f(z) = \frac{1}{(z - z_0)^m} \sum_{n=0}^{\infty} c_{n-m}(z - z_0)^n$$

$$= \frac{c_{-m}}{(z - z_0)^m} + \ldots + \frac{c_{-1}}{z - z_0} + \sum_{n=0}^{\infty} c_n(z - z_0)^n.$$

Now suppose that in the punctured neighbourhood of z_0 the function $f(z)$ is expanded by (26.112). This implies that in this neighbourhood $f(z)$ is analytic together with the function

$$g(z) = (z - z_0)^m f(z).$$

We can expand $g(z)$ by

$$g(z) = c_{-m} + c_{-m+1}(z - z_0) + \ldots.$$

It is seen that z_0 is a removable singularity of $g(z)$, and that there exists

$$\lim_{z \to z_0} g(z) = c_{-m} \neq 0.$$

Then

$$f(z) = \frac{g(z)}{(z - z_0)^m}$$

tends to ∞ as $z \to z_0$, i.e., z_0 is a pole of $f(z)$. ►

The point z_0 is a pole of $f(z)$ if and only if $g(z) = 1/f(z)$ is analytic in a neighbourhood of z_0 and $g(z_0) = 0$*).

The *order of a pole* of $f(z)$ is the order of the zero of $1/f(z)$.

Theorems 26.16 and 26.18 lead to the following assertion:

Theorem 26.19. *An isolated singularity is an essential singularity if and only if the principal part of the Laurent expansion in the punctured neighbourhood of the point contains an infinite number of nonzero terms.*

Examples. (1) The point $z_0 = 0$ is singularity of the function $f(z) = (e^z - 1)/z$.

◄ We have

$$\lim_{z \to 0} \frac{e^z - 1}{z} = 1.$$

Accordingly, $z_0 = 0$ is a removable singularity. The Laurent expansion of $f(z)$ about a zero point contains only the regular part

$$f(z) = \frac{1}{z}\left(1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots - 1\right)$$
$$= \frac{1}{1!} + \frac{z}{2!} + \frac{z^2}{3!} + \frac{z^3}{4!} + \ldots \quad ►$$

(2) $f(z) = \dfrac{1 - \cos z}{z^7}$.

◄ The function has a singularity $z_0 = 0$. We have

$$\lim_{z \to 0} \frac{1 - \cos z}{z^7} = \infty.$$

Consequently, $z_0 = 0$ is a pole. We expand $\cos z$ into a Taylor series in powers of z. We will then obtain the Laurent expansion of $f(z)$ about zero:

$$f(z) = \frac{1}{z^7}\left(\frac{z^2}{2!} - \frac{z^4}{4!} + \frac{z^6}{6!} - \ldots\right)$$
$$= \frac{1}{2!z^5} - \frac{1}{4!z^3} + \frac{1}{6!z} - \frac{z}{8!} + \frac{z^3}{10!} - \ldots$$

*) The function $g(z) = 1/f(z)$ can be made analytic by putting $g(z_0) = 0$.

The Laurent expansion of $f(z)$ about $z_0 = 0$ contains a finite number of terms with negative powers of z, namely three. Since the largest power of z in the denominator of the principal part of the expansion is five, then $z_0 = 0$ is a pole of order five. ►

The behaviour of a function about an essential singularity is characterized by the following

Theorem 26.20 (Sokhotsky theorem). *If z_0 is an essential singularity of a function $f(z)$, then for any complex number A (finite or infinite) there exists the sequence of points $z_k \to z_0$, such that*

$$\lim_{k\to\infty} f(z_k) = A.$$

Example. $f(z) = e^{1/z^2}$.

◄ The singularity here is $z_0 = 0$.

Let us consider the behaviour of the function on the real and imaginary axes: on the real axis $z = x$ and $f(x) = e^{1/x^2} \to +\infty$ as $x \to 0$, on the imaginary axis $z = iy$ and $f(iy) = e^{-1/y^2} \to 0$ as $y \to 0$.

And so there exists no limit of $f(z)$ as $z \to 0$, either finite or infinite. Hence, $z_0 = 0$ is an essential singularity of $f(z)$.

We now find the Laurent expansion of $f(z)$ in a neighbourhood of the zero point.

For any complex ζ we have

$$e^{\zeta} = 1 + \frac{\zeta}{1!} + \frac{\zeta^2}{2!} + \frac{\zeta^3}{3!} + \ldots.$$

We put $\zeta = 1/z^2$ to obtain

$$e^{1/z^2} = 1 + \frac{1}{1!z^2} + \frac{1}{2!z^4} + \frac{1}{3!z^6} + \ldots, \qquad |z| > 0.$$

The Laurent expansion contains an infinite number of terms with negative powers of z. ►

26.6 Residues. Basic Theorem of Residues. Application of Residues to Integrals

Residues. The *residue* of a function $f(z)$ at an isolated singularity z_0 is defined as a number

$$\operatorname{res} f(z_0) = \frac{1}{2\pi i} \int_{\gamma} f(\zeta)d\zeta, \tag{26.113}$$

where γ is a sufficiently small circumference $|z - z_0| = r$, i.e., on the circle $|z - z_0| \leqslant r$ there are no other singularities of $f(z)$.

From the formula for the coefficients of a Laurent expansion it follows directly that

$$\operatorname{res} f(z_0) = \frac{1}{2\pi i} \int_\gamma f(\zeta)d\zeta = c_{-1}. \tag{26.114}$$

The residue of $f(z)$ at an isolated singularity z_0 is thus equal to the coefficient at $(z - z_0)$ to the power minus one in the Laurent expansion of the function at z_0.

This suggests, in particular, that the residue at a removable singularity is equal to zero.

We will now provide some formulas to compute the residue at a pole of a function $f(z)$.

(1) z_0 is a pole of order one:

$$f(z) = \frac{c_{-1}}{z - z_0} + \sum_{n=0}^{\infty} c_n(z - z_0)^n.$$

Multiplying both sides of the relation by $z - z_0$ and passing to the limit as $z \to z_0$ gives

$$\lim_{z \to z_0} (z - z_0)f(z) = c_{-1}. \tag{26.115}$$

If $f(z)$ can be represented as a fraction

$$f(z) = \frac{\varphi(z)}{\psi(z)},$$

where $\varphi(z)$ and $\psi(z)$ are analytic functions, such that $\varphi(z_0) \neq 0$, $\psi(z_0) = 0$, $\psi'(z_0) \neq 0$, i.e., z_0 is a simple pole, then it follows from (26.115) that

$$c_{-1} = \lim_{z \to z_0} (z - z_0)\frac{\varphi(z)}{\psi(z)} = \lim_{z \to z_0} \frac{\varphi(z)}{\dfrac{\psi(z) - \psi(z_0)}{z - z_0}} = \frac{\varphi(z_0)}{\psi'(z_0)}.$$

Example. Let $f(z) = z/(z^2 + 1)$.

◄ The singularities $z = \pm i$ of $f(z)$ are simple poles. Therefore,

$$\operatorname{res} f(\pm i) = \frac{1}{\pm 2i} = \mp \frac{i}{2}. \quad ►$$

(2) z_0 is a pole or order m:

$$f(z) = \frac{c_{-m}}{(z - z_0)^m} + \dots + \frac{c_{-1}}{z - z_0} + \sum_{n=0}^{\infty} c_n(z - z_0)^n, \quad c_{-m} \neq 0.$$

To remove the negative powers of $z - z_0$ we multiply both sides of this by $(z - z_0)^m$

$$f(z)(z - z_0)^m = c_{-m} + \ldots + c_{-1}(z - z_0)^{m-1} + \sum_{n=0}^{\infty} c_n(z - z_0)^{n+m}.$$

Differentiating the resultant relation $m - 1$ times and passing to the limit as $z \to z_0$, we obtain

$$\operatorname{res} f(z_0) = c_{-1} = \frac{1}{(m-1)!} \lim_{z \to z_0} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)]. \tag{26.116}$$

Example. $f(z) = \dfrac{1}{(z^2 + 1)^2}$.

◄ The function has singularities at $z = \pm i$. These are poles of the second order. We will work out, say, $\operatorname{res} f(i)$. We have

$$\operatorname{res} f(i) = \frac{1}{1!} \lim_{z \to i} \frac{d}{dz} \left[(z - i)^2 \frac{1}{(z^2 + 1)^2} \right]$$
$$= \lim_{z \to i} \frac{d}{dz} \frac{1}{(z + i)^2} = -\lim_{z \to i} \frac{2}{(z + i)^3} = -\frac{2}{(2i)^3} = \frac{i}{4}. \quad ►$$

Theorem 26.21. *Let a function $f(z)$ be analytic throughout a domain D save for a finite number of isolated singularities $z_1, \ldots, z_n$. Then for any closed domain $\bar{G}$ lying in D and containing the points $z_1, \ldots, z_n$ we have*

$$\int_{\partial G} f(\zeta)\, d\zeta = 2\pi i \sum_{k=1}^{n} \operatorname{res} f(z_k). \tag{26.117}$$

◄ The theorem follows from the Cauchy theorem for a multiply connected domain. We construct the circumferences (Fig. 26.29)

$$\gamma_k: \ |z - z_k| = r \qquad (k = 1, \ldots, n)$$

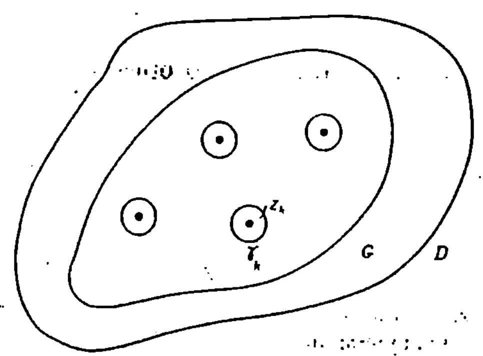


Fig. 26.29

of radius r that is so small that the circles $U_k = |z - z_k| \leqslant r$, bounded by them lie in G and do not overlap.

We denote by G^* the domain that remains when we remove from G the circles $U_1, \ldots, U_n$. The function $f(z)$ is analytic in G^* and continuous in its closure $\overline{G}^*$. Therefore, by the Cauchy theorem for a multiply connected domain we have

$$\int_{\partial G} f(\zeta)\, d\zeta = \sum_{k=1}^{n} \int_{\gamma_k} f(\zeta)\, d\zeta.$$

From this, using the definition of the residue

$$\int_{\gamma_k} f(\zeta)\, d\zeta = 2\pi i f(z_k),$$

we arrive at the relation (26.117) required. ►

Residue at an infinitely removed point. We say that a function $f(z)$ is analytic at $z = \infty$, if the function

$$g(\zeta) = f\left(\frac{1}{\zeta}\right)$$

is analytic at the point $\zeta = 0$*).

For example, the function

$$f(z) = \sin\frac{1}{z}$$

is analytic at $z = \infty$, since the function

$$g(\zeta) = f\left(\frac{1}{\zeta}\right) = \sin\zeta$$

is analytic at $\zeta = 0$.

Let a function $f(z)$ be analytic in some neighbourhood of ∞ (save for the point $z = \infty$ itself).

The point $z = \infty$ is called an *isolated singularity* of $f(z)$, if in some neighbourhood of the point there are no other singularities of $f(z)$.

The function

$$f(z) = \frac{1}{\sin z}$$

has at infinity a nonisolated singularity: the poles $z_k = k\pi$ of the function accumulate at infinity if $k \to \infty$.

*) This shoud be taken to mean that a function $g(\zeta) = f(1/\zeta)$ can be made analytic by putting $g(0) = \lim\limits_{z\to\infty} f(z)$.

We say that $z = \infty$ is a *removable singularity*, a *pole* or an *essential singularity* of $f(z)$ depending on whether $\lim\limits_{z \to \infty} f(z)$ is finite, infinite or nonexistent.

The type of an infinite point in terms of the Laurent expansion is determined by other criteria than for finite singularities.

Theorem 26.22. *If $z = \infty$ is a removable singularity of a function $f(z)$, then the Laurent expansion of $f(z)$ in a neighbourhood of the point contains no positive powers of z. If $z = \infty$ is a pole, then this expansion contains a finite number of positive powers of z, and if the singularity is an essential one, an infinite number of positive powers of z.*

We define the Laurent expansion of $f(z)$ in a neighbourhood of an infinite point to be an expansion into a Laurent series such that it converges everywhere outside a circle of a sufficiently large radius R with centre at $z = 0$ (except perhaps for $z = \infty$).

Let a function $f(z)$ be analytic in some neighbourhood of $z = \infty$ (save perhaps for the point itself).

The *residue of $f(z)$ at infinity* is the quantity

$$\operatorname{res} f(\infty) = \frac{1}{2\pi i} \int\limits_{\gamma^-} f(\zeta)\, d\zeta, \tag{26.118}$$

where γ^- is a sufficiently large circumference $|z| = \varrho$ traced clockwise (so that the neighbourhood of $z = \infty$ will lie to the left, as in the case of the finite point $z = z_0$).

It follows from the definition that the residue of the function at infinity is equal to the coefficient at z^{-1} in the Laurent expansion of $f(z)$ about $z = \infty$ taken with the opposite sign

$$\operatorname{res} f(\infty) = -c_{-1}. \tag{26.119}$$

Example. Consider the function $f(z) = (z + 1)/z$.

◄ We have $f(z) = 1 + 1/z$. This expression can be viewed as the Laurent expansion of the function in the neighbourhood of $z = \infty$. It can easily be seen that

$$\lim_{z \to \infty} f(z) = 1,$$

so that the point $z = \infty$ is a removable singularity, and we put, as usual, $f(\infty) = 1$. Here $c_{-1} = 1$ and hence

$$\operatorname{res} f(\infty) = -1. ►$$

This example suggests that the residue of an analytic function about a removable singularity at infinity (unlike a finite removable singularity) may be nonzero.

The well-known Taylor expansions of the functions e^z, $\cos z$, $\sin z$, $\cosh z$, and $\sinh z$ can also be regarded as Laurent expansions in the neighbourhood of $z = \infty$. Since all of these expansions contain an infinite number of positive powers of z, the above functions have an essential singularity at $z = \infty$.

Theorem 26.23. *If a function $f(z)$ has on the extended complex plane a finite number of singularities, then the sum of all its residues, including the one at infinity, will be zero.*

Therefore, if $z_1, \ldots, z_n$ are finite singularities of $f(z)$, then

$$\operatorname{res} f(\infty) + \sum_{k=1}^{n} \operatorname{res} f(z_k) = 0 \tag{26.120}$$

or

$$\operatorname{res} f(\infty) = -\sum_{k=1}^{n} \operatorname{res} f(z_k). \tag{26.121}$$

This relation can conveniently be used to compute some integrals.

Example. Take the integral $I = \int\limits_{|z|=2} \frac{dz}{1+z^8}$.

◄ The poles (finite) of the integrand $f(z) = 1/(1 + z^8)$ are the roots z_k of the equation $z^8 = -1$, all of which lie inside the circle $|z| = 2$. In the neighbourhood of the point $z = \infty$ the function $f(z)$ has the following expansion

$$f(z) = \frac{1}{z^8} \frac{1}{1 + \frac{1}{z^8}} = \frac{1}{z^8} - \frac{1}{z^{16}} + \frac{1}{z^{24}} - \ldots$$

from which it is seen that $\operatorname{res} f(\infty) = -c_{-1} = 0$.

By Theorem 26.23,

$$I = 2\pi i \sum_{k=1}^{\infty} \operatorname{res} f(z_k) = -2\pi i \operatorname{res} f(\infty) = 0. ►$$

Application of residues in computing definite integrals. Residues are widely used to take definite integrals, notably integrals of *rational function.*

Theorem 26.24. *Let $f(x)$ be a rational function, i.e.,*

$$f(x) = \frac{P_n(x)}{Q_m(x)},$$

where $P_n(x)$ and $Q_m(x)$ are polynomials of degrees n and m, respectively. If $f(x)$ is continuous on the entire real axis ($Q_m(x) \neq 0$) and $m \geqslant n + 2$, i.e., the degree of the denominator is, at least, by two larger than the degree

of the numerator, then

$$\int_{-\infty}^{\infty} f(x)\,dx = 2\pi i\sigma,$$

where σ is the sum of the residues of the function $f(z)$ at all poles lying in the upper half-plane.

◄ Consider the closed contour γ made up of the segment of the real axis $-R \leqslant x \leqslant R$ and the upper half-circle γ_R: $|z| = R$, $\operatorname{Im} z > 0$. We can take R so large that the inside of the domain bounded by γ will contain all the poles $z_1, \ldots, z_l$ of $f(z)$ that lie in the upper half-plane.

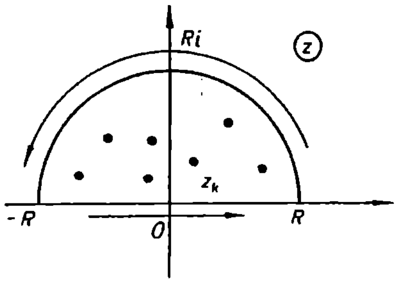


Fig. 26.30

By the basic theorem on residues (Fig. 26.30),

$$\int_{\gamma} f(\zeta)\,d\zeta = \int_{-R}^{R} f(x)\,dx + \int_{\gamma_R} f(\zeta)\,d\zeta = 2\pi i \sum_{k=1}^{l} \operatorname{res} f(z_k).$$

We estimate $\int\limits_{\gamma_R}$. By the condition imposed on the degree of $P_n(z)$ and $Q_m(z)$ there are positive numbers R_0 and M such that for $|z| > R_0$

$$|f(z)| < \frac{M}{|z|^2}.$$

By Property (6) of integrals of a function of a complex variable we have for $R > R_0$

$$\left| \int_{\gamma_R} f(\zeta)\,d\zeta \right| \leqslant \int_{\gamma_R} |f(\zeta)|\,|d\zeta| < \frac{M}{R^2}\pi R = \frac{\pi M}{R} \to 0$$

as $R \to \infty$.

We pass in

$$\int_{-R}^{R} f(x)\,dx + \int_{\gamma_R} f(\zeta)\,d\zeta = 2\pi i \sum_{k=1}^{l} \operatorname{res} f(z_k)$$

to the limit as $R \to \infty$. Notice that the right-hand side of this is independent of R, and the second term on the left-hand side tends to zero. It follows that the limit of the first term exists and is

$$\int_{-\infty}^{\infty} f(x)\,dx = 2\pi i \sum_{k=1}^{l} \operatorname{res} f(z_k), \tag{26.122}$$

where $z_1, \ldots, z_l$ are all poles of $f(z)$ lying in the upper half-plane. ►

Example. Take the integral $I = \int_0^{\infty} \dfrac{x^2\,dx}{(x^2 + a^2)^2}$, $a > 0$.

◄ Since the integrand

$$f(x) = \frac{x^2}{(x^2 + a^2)^2}$$

is an even function, we have

$$I = \frac{1}{2} \int_{-\infty}^{\infty} \frac{x^2\,dx}{(x^2 + a^2)^2}.$$

Consider the function

$$f(z) = \frac{z^2}{(z^2 + a^2)^2},$$

which is coincident with $f(x)$ on the real axis, i.e., at $z = x$. In the upper half-plane $f(z)$ has an isolated singularity $z = ai$, which is a pole of the second order. The residue of $f(z)$ at $z = ai$ will be

$$\operatorname{res} f(ai) = \lim_{z \to ai} \frac{d}{dz} [f(z)(z - ai)^2]$$

$$= \lim_{z \to ai} \frac{d}{dz} \left[\frac{z^2}{(z + ai)^2} \right] = \lim_{z \to ai} \frac{2aiz}{(z + ai)^3} = \frac{1}{4ai}.$$

Using (26.122) we find

$$I = \frac{1}{2} \int_{-\infty}^{\infty} \frac{x^2\,dx}{(x^2 + a^2)^2} = \frac{1}{2} 2\pi i \frac{1}{4ai} = \frac{\pi}{4a}. \blacktriangleright$$

Consider the integral of the form

$$I = \int_0^{2\pi} R(\cos x, \sin x)\,dx,$$

where $R(u, v)$ is a rational function of u and v. We introduce the complex variable $z = e^{ix}$. Then

$$dx = \frac{dz}{iz}, \qquad \cos x = \frac{z^2+1}{2z}, \qquad \sin x = \frac{z^2-1}{2iz}. \tag{26.123}$$

Clearly, here $|z| = 1$, $0 \leqslant x \leqslant 2\pi$. The original integral thus becomes an integral of a function of a complex variable round a closed contour

$$I = \int\limits_{\gamma} R\left(\frac{z^2+1}{2z}, \quad \frac{z^2-1}{2iz}\right)\frac{dz}{iz} = \int\limits_{\gamma} F(z)\,dz,$$

where γ is the circumference of unit radius with centre at the origin: $|z| = 1$.

According to the basic theorem on residues, the resultant integral is equal to $2\pi i\sigma$, where σ is the sum of the residues of the integrand $F(z)$ at poles lying inside γ.

Example. Take the integral

$$I = \int\limits_0^{2\pi} \frac{dx}{(a + b\cos x)^2}, \qquad a > b > 0.$$

◄ Substituting $z = e^{ix}$ and rearranging (see (26.123)) gives

$$I = \frac{4}{i}\int\limits_{\gamma} \frac{z\,dz}{(bz^2 + 2az + b)^2} = \frac{4}{i}\,2\pi i \sum_{k=1}^{n} \operatorname{res} F(z_k).$$

Inside the unit circle $(a > b > 0)$ lies only one pole (of the second order)

$$z_1 = \frac{-a + \sqrt{a^2 - b^2}}{b} \qquad (n = 1).$$

The residue of the function

$$F(z) = \frac{z}{(bz^2 + 2az + b)^2}$$

at the point z_1 is

$$\operatorname{res} F(z_1) = \lim_{z \to z_1} \frac{d}{dz}\left[\frac{(z - z_1)^2 z}{b^2(z - z_1)^2(z - z_2)^2}\right] = \frac{a}{4(a^2 - b^2)^{3/2}}.$$

Thus

$$I = \frac{2\pi a}{(a^2 - b^2)^{3/2}}.$$ ►

We next consider *integrals of the type*

$$\int_0^{\infty} R(x)\cos ax\,dx, \qquad \int_0^{\infty} R(x)\sin ax\,dx,$$

where $R(x)$ is a proper rational fraction, $a > 0$ is a real number.

When handling such integrals the following lemma is often helpful:

Jordan's lemma. *Let a function $f(z)$ be analytic in the upper half-plane* $\operatorname{Im} z > 0$, *except for a finite number of isolated singularities, and for* $|z| \to \infty$ *it tends to zero uniformly in* $\arg z$. *Then for any positive a*

$$\lim_{R\to\infty} \int_{\gamma_R} f(z)\,e^{iaz}\,dz = 0 \tag{26.124}$$

where γ_R is the upper semi-circumference $|z| = R$, $\operatorname{Im} z > 0$.

◄ The condition for $f(z)$ to tend uniformly to zero suggests that on the semi-circumference γ_R

$$|f(z)| \leqslant M_R,$$

where $M_R \to 0$ as $R \to \infty$.

We will estimate the integral under consideration. We notice that on γ_R

$$|e^{iaz}| = |e^{iaR(\cos\varphi + i\sin\varphi)}| = |e^{-aR\sin\varphi + iaR\cos\varphi}| = e^{-aR\sin\varphi},$$

$$dz = d(R\,e^{i\varphi}) = iR\,e^{i\varphi}\,d\varphi.$$

Hence, $|dz| = R\,d\varphi$. We get

$$\left|\int_{\gamma_R} f(z)\,e^{iaz}\,dz\right| \leqslant \int_{\gamma_R} |f(z)|\,|e^{iaz}|\,|dz|$$

$$\leqslant M_R \cdot R \int_0^{\pi} e^{-aR\sin\varphi}\,d\varphi = 2M_R \cdot R \int_0^{\frac{\pi}{2}} e^{-aR\sin\varphi}\,d\varphi. \tag{26.125}$$

By the well-known inequality (Fig. 26.31) $\sin\varphi \geqslant 2\varphi/\pi$, which holds for $0 < \varphi \leqslant \pi/2$. To prove this, it is sufficient to note that

$$\left(\frac{\sin\varphi}{\varphi}\right)' = \frac{\cos\varphi}{\varphi^2}(\varphi - \tan\varphi) < 0 \qquad \text{for} \qquad 0 < \varphi < \frac{\pi}{2},$$

and so the function $(\sin\varphi)/\varphi$ decreases on the half-interval $(0, \pi/2]$ and

$$\int_0^{\pi/2} e^{-aR\sin\varphi}\,d\varphi \leqslant \int_0^{\pi/2} e^{\frac{-2aR}{\pi}\varphi}\,d\varphi = \frac{\pi}{2aR}(1 - e^{-aR}). \tag{26.126}$$

Comparing (26.125) and (26.126) gives

$$\left| \int_{\gamma_R} f(z)\, e^{iaz}\, dz \right| \leqslant \frac{\pi M_R}{a} (1 - e^{-aR}) \to 0$$

as $R \to \infty$. ►

Example. Compute the integral

$$I = \int_0^{\infty} \frac{x \sin ax}{x^2 + k^2}\, dx, \qquad a > 0, \qquad k > 0.$$

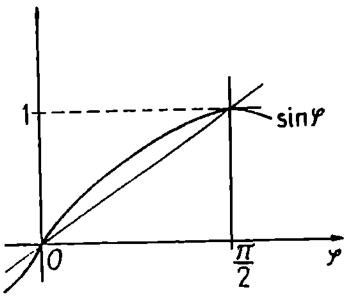


Fig. 26.31

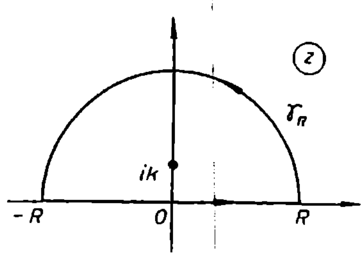


Fig. 26.32

◄ We introduce the auxiliary function

$$h(z) = \frac{z\, e^{iaz}}{z^2 + k^2}$$

It is easily seen that if $z = x$, then $\operatorname{Im} h(z)$ coincides with the integrand (Fig. 26.32)

$$f(x) = \frac{x \sin ax}{x^2 + k^2}.$$

Consider the contour indicated on the figure. At sufficiently large R the function

$$g(z) = \frac{z}{z^2 + k^2},$$

according to the relation $|z^2 + k^2| \geqslant |z|^2 - k^2$, obeys on the arc γ_R the inequality

$$|g(z)| < \frac{R}{R^2 - k^2} \to 0$$

as $R \to \infty$. Hence, by the Jordan lemma

$$\lim_{R\to\infty} \int_{\gamma_R} \frac{z\,e^{iaz}}{z^2+k^2}\,dz = 0. \tag{26.127}$$

For any $R > k$, according to the basic theorem on residues, we have

$$\int_{-R}^{R} \frac{x\,e^{iax}}{x^2+k^2}\,dx + \int_{\gamma_R} \frac{z\,e^{iaz}}{z^2+k^2}\,dz = 2\pi i\sigma, \tag{26.128}$$

where

$$\sigma = \operatorname*{res}_{z=ik}\left[\frac{z\,e^{iaz}}{z^2+k^2}\right] = \lim_{z\to ik} \frac{z\,e^{iaz}(z-ik)}{z^2+k^2} = \frac{e^{-ak}}{2}.$$

Passing to the limit in (26.128) and taking (26.127) into account, we find

$$\int_{-\infty}^{\infty} \frac{x\,e^{iax}}{x^2+k^2}\,dx = \pi i\,e^{-ak}.$$

Separating the real and imaginary parts on the left and right we will have

$$\int_{-\infty}^{\infty} \frac{x\sin ax}{x^2+k^2}\,dx = \pi e^{-ak}.$$

Because the integrand $f(z)$ is an even function, we find

$$I = \frac{\pi}{2}e^{-ak}. ►$$

In this example the function $f(z)$ has no singularities on the real axis. But some additional measures enable us to apply the method even when $f(z)$ has singularities (simple poles) on the real axis. We will illustrate this by an example.

Example. Compute the integral

$$I = \int_{0}^{\infty} \frac{\sin ax}{x(x^2+b^2)}\,dx, \qquad a > 0, \qquad b > 0. \tag{26.129}$$

◄ The function

$$h(z) = \frac{e^{iaz}}{z(z^2+b^2)}$$

has the following properties:

(1) $\operatorname{Im} h(z)$ at $z = x$ coincides with the integrand;

(2) $h(z)$ has a singularity on the real axis (a simple pole at $z = 0$).

Consider in the upper half-plane $\operatorname{Im} z \geqslant 0$ a closed contour Γ consisting of segments of the real axis $[-R, -r]$, $[r, R]$ and arcs of semicircumferences γ_r, $|z| = r$, and γ_R, $|z| = R$ (Fig. 26.33).

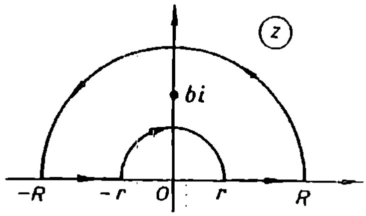


Fig. 26.33

Inside the contour lies only one pole of $h(z)$, i.e., the point $z = bi$. By the basic theorem on residues

$$\int_{\Gamma} h(z)\,dz = \int_{-R}^{-r} \frac{e^{iax}}{x(x^2+b^2)}\,dx + \int_{\gamma_R} \frac{e^{iaz}}{z(z^2+b^2)}\,dz$$
$$+ \int_{r}^{R} \frac{e^{iax}}{x(x^2+b^2)}\,dx + \int_{\gamma_R} \frac{e^{iaz}}{z(z^2+b^2)}\,dz = 2\pi i\sigma, \tag{26.130}$$

where

$$\sigma = \operatorname*{res}_{z=bi} \frac{e^{iaz}}{z(z^2+b^2)} = \lim_{z\to bi} \frac{e^{iaz}(z-bi)}{z(z^2+b^2)} = -\frac{e^{-ab}}{2b^2}.$$

We rearrange at first the sum of the integrals along the segments $[-R, -r]$ and $[r, R]$ on the real axis. Replacing in the first term on the right of (26.130) x by $-x$ and combining it with the third term gives

$$\int_{-R}^{-r} \frac{e^{iax}}{x(x^2+b^2)}\,dx + \int_{r}^{R} \frac{e^{iax}}{x(x^2+b^2)}\,dx$$
$$= \int_{r}^{R} \frac{e^{iax}-e^{-iax}}{x(x^2+b^2)}\,dx = 2i\int_{r}^{R} \frac{\sin ax\,dx}{x(x^2+b^2)}.$$

We now turn to the second term in (26.130). Since

$$\lim_{z\to 0} \frac{e^{iaz}}{z^2+b^2} = \frac{1}{b^2},$$

i.e.,

$$\frac{e^{iaz}}{z^2+b^2} = \frac{1}{b^2} + g(z),$$

where $\lim\limits_{z\to 0} g(z) = 0$, then the integrand $h(z)$ can be represented as follows:

$$h(z) = \frac{1}{b^2 z} + \frac{g(z)}{z}.$$

Then

$$\int\limits_{\gamma_r} h(z)\,dz = \frac{1}{b^2}\int\limits_{\gamma_r} \frac{dz}{z} + \int\limits_{\gamma_r} \frac{g(z)}{z}\,dz.$$

Putting $z = re^{i\varphi}$, we find

$$\frac{1}{b^2}\int\limits_{\gamma_r} \frac{dz}{z} = \frac{1}{b^2}\int\limits_{\pi}^{0} \frac{ire^{i\varphi}}{re^{i\varphi}}\,d\varphi = -\frac{i\pi}{b^2},$$

$$\int\limits_{\gamma_r} \frac{g(z)}{z}\,dz = i\int\limits_{\pi}^{0} g(re^{i\varphi})\,d\varphi \to 0, \qquad r\to 0.$$

By the Jordan lemma, the fourth term in (26.130) as $R\to\infty$ tends to zero, because, the function $f(z) = \dfrac{1}{z(z^2+b^2)}$ tends to zero as $|z|\to\infty$:

$$\lim_{R\to\infty} \int\limits_{\gamma_R} \frac{e^{iaz}}{z(z^2+b^2)}\,dz = 0. \tag{26.131}$$

And so, as $R\to\infty$ and $r\to 0$, (26.130) becomes

$$2i\int\limits_{0}^{\infty} \frac{\sin ax\,dx}{x(x^2+b^2)} - \frac{\pi i}{b^2} = -\pi i\,\frac{e^{-ab}}{b^2}.$$

Hence

$$\int_0^\infty \frac{\sin ax\,dx}{x(x^2+b^2)} = \frac{\pi}{2b^2}(1-e^{-ab}). \blacktriangleright$$

Consider the *Fresnel integral*

$$\int_0^\infty \cos x^2\,dx, \qquad \int_0^\infty \sin x^2\,dx.$$

We introduce the auxiliary function $f(z) = e^{iz^2}$. Inside the contour Γ shown in Fig. 26.34 ($OA = OB = r$, $\angle AOB = \pi/4$) the function is analytic, and by the Cauchy theorem

$$\int_\Gamma e^{iz^2}\,dz = \int_{OA} e^{ix^2}\,dx + \int_{\gamma_r} e^{iz^2}\,dz + \int_{BO} e^{iz^2}\,dz = 0. \qquad (26.132)$$

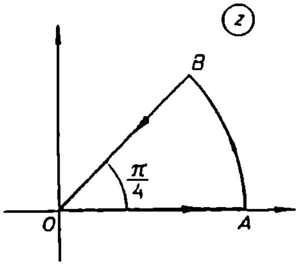


Fig. 26.34

We show that

$$\lim_{r\to\infty} \int_{\gamma_r} e^{iz^2}\,dz = 0.$$

Putting $z^2 = \zeta$, we obtain $dz = \dfrac{d\zeta}{2\sqrt{\zeta}}$ and

$$\int_{\gamma_r} e^{iz^2}\,dz = \int_{\Gamma_{r^2}} \frac{e^{i\zeta}}{2\sqrt{\zeta}}\,d\zeta,$$

where Γ_{r^2} is a semicircle of radius r^2.

The function $g(\zeta) = 1/(2\sqrt{\zeta})$ meets the conditions of the Jordan lemma, and hence

$$\lim_{r\to\infty} \int_{\Gamma_{r_2}} \frac{e^{i\zeta}}{2\sqrt{\zeta}}\, d\zeta = \lim_{r\to\infty} \int_{\gamma_r} e^{iz^2}\, dz = 0.$$

On the segment BO: $z = \varrho\, e^{i\pi/4}$, $z^2 = \varrho^2 e^{i\pi/2} = \varrho^2 i$, $0 \leqslant \varrho \leqslant r$. Hence

$$\int_{BO} e^{iz^2}\, dz = \int_r^0 e^{-\varrho^2} e^{i\frac{\pi}{4}}\, d\varrho = -e^{i\frac{\pi}{4}} \int_0^r e^{-\varrho^2}\, d\varrho.$$

Passing in (26.132) to the limit as $r \to \infty$, gives

$$\int_0^\infty e^{ix^2}\, dx = e^{i\frac{\pi}{4}} \int_0^\infty e^{-\varrho^2}\, d\varrho. \tag{26.133}$$

Knowing that,

$$\int_0^\infty e^{-\varrho^2}\, d\varrho = \frac{\sqrt{\pi}}{2},$$

we have

$$\int_0^\infty \cos x^2\, dx + i \int_0^\infty \sin x^2\, dx = \frac{1}{2}\sqrt{\frac{\pi}{2}} + \frac{i}{2}\sqrt{\frac{\pi}{2}}.$$

Hence

$$\int_0^\infty \cos x^2\, dx = \frac{1}{2}\sqrt{\frac{\pi}{2}}, \qquad \int_0^\infty \sin x^2\, dx = \frac{1}{2}\sqrt{\frac{\pi}{2}}. \blacktriangleright$$

Exercises

Find the real and imaginary parts of the function:

1. $w = \bar{z} - iz^2$. 2. $w = z^2 + i$. 3. $w = \frac{1}{z}$. 4. $w = \frac{\bar{z}}{z}$.

Find the images of the real and imaginary axes for the mappings:

5. $w = \frac{z+1}{z-1}$. 6. $w = 1 + \frac{1}{z}$.

Prove that the function is continuous in the entire complex plane:

7. $w = \bar{z}$. 8. $w = \operatorname{Re} z$. 9. $w = \operatorname{Im} z$.

Using the Cauchy-Riemann conditions, find whether or not the function is analytic at least at one point:

10. $z^2\bar{z}$. 11. $ze^{\bar{z}}$. 12. $|z|\,\bar{z}$. 13. $|z|\operatorname{Re} z$. 14. e^{z^2}.

Restore a function $f(z)$ that is analytic in a neighbourhood of the point z_0 from the known real part $u(x, y)$ or from the known imaginary part $v(x, y)$ and the value $f(z_0)$:

15. $u = \dfrac{x}{x^2 + y^2}$, $f(\pi) = \dfrac{1}{\pi}$. 16. $v = 2xy + 2y$, $f(i) = 2i - 1$.

Show that the following functions are harmonic:

17. $x^2 + 2x - y^2$. 18. $2e^x \cos y$. 19. $\tan^{-1}\dfrac{y}{x}$. 20. $\ln(x^2 + y^2)$.

Show whether or not the following functions can be real or imaginary parts of the analytic function $f(z) = u(x, y) + iv(x, y)$:

21. $u = x^2 - y^2 + 2xy$. 22. $u = x^2$. 23. $v = \dfrac{x^2 + 1}{2}y$.

Find the real and imaginary parts of the following functions:
24. $w = e^{-z}$. 25. $w = \sin z$. 26. $w = \cosh(z - i)$. 27. $w = \tan z$.

Find the modulus and principal value of the argument of the function at a point z_0:

28. $w = \cos z$, $z_0 = \dfrac{\pi}{2} + i\ln 2$. 29. $w = \sinh z$, $z_0 = 1 + i\dfrac{\pi}{2}$. 30. $w = ze^z$, $z_0 = i\pi$.

Find the logarithm of the following numbers:
31. e. 32. $-i$. 33. i. 34. $-1 - i$.

Solve the equations:
35. $e^{-z} + 1 = 0$; 36. $e^z + i = 0$. 37. $\sin z = \pi i$. 38. Take the integral $\int\limits_\gamma (1 + i - 2\bar{z})\,dz$, where γ is a line connecting the points $z_1 = 0$ and $z_2 = 1 + i$, which is: (a) a segment; (b) an arc of the parabola $y = x^2$; (c) a broken line $z_1z_3z_2$, $z_3 = 1$.

39. Take the integral $\int\limits_\gamma (z^2 + z\bar{z})\,dz$, where γ is the semi-circumference $|z| = 1$, $0 \leqslant \arg z \leqslant \pi$.

Take the integrals:

40. $\displaystyle\int_{1+i}^{-1-i} (2z + 1)\,dz$. 41. $\displaystyle\int_0^{1+i} e^z\,dz$. 42. $\displaystyle\int_{-i}^{i} z e^{z^2}\,dz$.

43. Take the integral $\displaystyle\int_\gamma \frac{dz}{\sqrt{z}}$, where γ is the upper half of the circumference $|z| = 1$ (take the branch of $\sqrt{z}$ for which $\sqrt{1} = 1$).

44. Take the integral $\displaystyle\int_\gamma \frac{\ln z}{z}\,dz$, where γ is the segment connecting the points $z_1 = 1$ and $z_2 = i$.

Take integrals:

45. $\int\limits_{|z|=1} \frac{e^z}{z^2 + 2z}\, dz$. 46. $\int\limits_{|z-i|=1} \frac{e^{iz}}{z^2 + 1}\, dz$.

47. $\int\limits_{|z|=1} \frac{\cos z}{z^3}\, dz$. 48. $\int\limits_{|z|=\frac{1}{2}} \frac{1 - \sin z}{z^2}\, dz$.

Find the radius of convergence for the series:

49. $\sum\limits_{n=1}^{\infty} e^{in} z^n$. 50. $\sum\limits_{n=1}^{\infty} i^n z^n$.

51. $\sum\limits_{n=1}^{\infty} \left(\frac{z}{in}\right)^n$. 52. $\sum\limits_{n=1}^{\infty} \sin\frac{\pi i}{n} z^n$.

53. $\sum\limits_{n=1}^{\infty} (1 + i)^n z^n$.

Expand the functions into a Taylor series and find the radius of convergence of the resultant series:

54. $\sin(2z + 1)$ in $z + 1$. 55. $\cos z$ in $z + \frac{\pi}{4}$.

56. $\frac{1}{3z + 1}$ in $z + 2$. 57. $\frac{z}{z^2 + 1}$ in z.

58. $\sinh^2 z$ in z.

Find the zeros of the functions and their orders:

59. $z^4 + 4z^2$. 60. $(\sin z)/z$. 61. $z^2 \sin z$. 62. $1 + \cosh z$.

Determine the interval of convergence for the series:

63. $\sum\limits_{n=1}^{\infty} \frac{e^n}{(iz)^n}$. 64. $\sum\limits_{n=1}^{\infty} \left(\frac{2}{z}\right)^n + \sum\limits_{n=0}^{\infty} \left(\frac{z}{4}\right)^n$.

65. $\sum\limits_{n=1}^{\infty} \frac{(-1)^n}{n^4 z^n} + \sum\limits_{n=0}^{\infty} \frac{z^n}{n 2^n}$. 66. $\frac{1}{z} + \sum\limits_{n=0}^{\infty} z^n$.

Find the Laurent expansion about the point $z = 0$ of the functions:

67. $\frac{\sin z}{z^2}$. 68. $\frac{e^z}{z^3}$. 69. $z^4 \cos\frac{1}{z}$. 70. $\frac{1 - e^{-z}}{z^3}$.

Find the Laurent expansion in the annulus:

71. $\frac{1}{z^2 + z}$, $0 < |z| < 1$. 72. $\frac{1}{z^2 + z}$, $1 < |z| < \infty$.

73. $\dfrac{2z+3}{z^2+3z+2}$, $1 < |z| < 2$.

74. $\dfrac{1}{(z^2-4)^2}$, $4 < |z+2| < \infty$.

Find the singularities of the following functions and determine their nature:

75. $\dfrac{1-\cos z}{z^2}$. 76. $e^{-1/z}$.

77. $\dfrac{\sin z}{z^2}$. 78. $\dfrac{z^2-1}{z^6+2z^5+z^4}$.

Find the residues of the functions at their singularities:

79. $\dfrac{e^z}{z^3(z-1)}$. 80. $\cos\dfrac{1}{z} + z^3$. 81. $z^2\sin\dfrac{1}{z}$. 82. $\dfrac{e^{\pi z}}{z-i}$.

Take the integrals:

83. $\displaystyle\int\limits_{|z|=2} \frac{e^z\,dz}{z^3(z+1)}$. 84. $\displaystyle\int\limits_{|z|=1} \frac{z^2\,dz}{\sin^3 z\cos z}$.

85. $\displaystyle\int\limits_{|z|=1} z^3\sin\frac{1}{z}\,dz$. 86. $\displaystyle\int\limits_{|z-1|=1} \frac{e^{2z}\,dz}{z^3-1}$.

Determine the nature of the point at infinity:

87. $\dfrac{z+1}{z^4}$. 88. $\dfrac{e^z}{z^2}$. 89. $\cos\dfrac{1}{z}$. 90. $z^3e^{1/z}$.

Take the integrals:

91. $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{(x^2+1)^3}$. 92. $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{1+x^6}$.

93. $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{(x^2+2x+2)^2}$. 94. $\displaystyle\int_{-\infty}^{\infty} \frac{\cos x\,dx}{x^2+9}$.

95. $\displaystyle\int_{0}^{\infty} \frac{x^2\cos x\,dx}{(x^2+1)^2}$. 96. $\displaystyle\int_{0}^{\infty} \frac{x^3\sin x\,dx}{(x^2+1)^2}$.

97. $\displaystyle\int_{0}^{\infty} \frac{\cos ax - \cos bx}{x^2}\,dx$, $a > 0$, $b > 0$.

98. $\displaystyle\int_0^{2\pi} \frac{dx}{1 - 2p\cos x + p^2}$, $0 < p < 1$.

99. $\displaystyle\int_0^{2\pi} \frac{\cos x\, dx}{1 - 2p\sin x + p^2}$, $0 < p < 1$.

Answers

1. $u = x + 2xy$, $v = y^2 - x^2 - y$. 2. $u = x^2 - y^2$, $v = 1 + 2xy$. 3. $u = \dfrac{x}{x^2 + y^2}$, $v = \dfrac{y}{x^2 + y^2}$. 4. $u = \dfrac{x^2 - y^2}{x^2 + y^2}$, $v = -\dfrac{2xy}{x^2 + y^2}$. 5. The x-axis goes into the u-axis; as x changes from $-\infty$ to $+\infty$ u changes from $+1$ to $-\infty$ and from $+\infty$ to $+1$ (the point $+1$ is excluded); the y-axis goes into the circle $u^2 + v^2 = 1$. 6. The x-axis goes into the u-axis as above; the y-axis goes into the line $u = 1$ that is traced from the point $+1$ to $1 + i\infty$ and from $1 - i\infty$ to 1 (the point $+1$ is excluded). 10. No. 11. Yes. 12. No. 13. No. 14. Yes. 15. $\dfrac{1}{z}$. 16. $z^2 + 2z$. 21. Yes. 22. No. 23. No. 24. $u = e^{-x}\cos y$, $v = e^{-x}\sin y$. 25. $u = \sin x\cosh y$, $v = \cos x\sinh y$. 26. $u = \cosh x\cos(y - 1)$, $v = \sinh x\sin(y - 1)$. 27. $u = \dfrac{\sin x\cos x}{\cosh^2 y - \sin^2 x}$, $v = \dfrac{\sinh y\cosh y}{\cosh^2 y - \sin^2 x}$. 28. $\varrho = \dfrac{3}{4}$, $\varphi_0 = -\dfrac{\pi}{2}$. 29. $\varrho = \cosh 1$, $\varphi_0 = \dfrac{\pi}{2}$. 30. $\varrho = \pi$, $\varphi_0 = -\dfrac{\pi}{2}$.

31. $1 + i2\pi k$. 32. $i\left(2k - \dfrac{1}{2}\right)\pi$. 33. $i\left(2k + \dfrac{1}{2}\right)\pi$. 34. $\ln\sqrt{2} + i\left(2k - \dfrac{3}{4}\right)\pi$ (everywhere $k = 0, \pm1, \pm2, \ldots$). 35. $z_k = i(2k + 1)\pi$. 36. $z_k = i\left(2k - \dfrac{1}{2}\pi\right)$. 37. $z_{2k} = 2\pi k - i\ln(\sqrt{\pi^2 + 1} - \pi)$, $z_{2k+1} = (2k + 1)\pi - i\ln(\sqrt{\pi^2 + 1} - \pi)$ (everywhere $k = 0, \pm1, \pm2$). 38. (a) $2(i - 1)$, (b) $-2 + \dfrac{4}{3}i$, (c) -2. 39. $-8/3$. 40. $-2 - 2i$. 41. $e\cos 1 - 1 + ie\sin 1$.

42. 0. 43. $-2 + 2i$. 44. $-\pi^2/8$. 45. πi. 46. π/e. 47. $-\pi i$. 48. $-2\pi i$. 49. 1. 50. 1. 51. ∞. 52. 1. 53. $\dfrac{1}{2}$.

54. $-\sin 1 + 2(z + 1)\cos 1 + \dfrac{2^2}{2!}(z + 1)^2\sin 1 - \dfrac{2^3}{3!}(z + 1)^3\cos 1 - \ldots$, $R = \infty$.

55. $\dfrac{1}{\sqrt{3}}\left[1 + \left(z + \dfrac{\pi}{4}\right) - \dfrac{1}{2!}\left(z + \dfrac{\pi}{4}\right)^2 - \dfrac{1}{3!}\left(z + \dfrac{\pi}{4}\right)^3 + \ldots\right]$, $R = \infty$

56. $-\dfrac{1}{5}\left[1 + \dfrac{3}{5}(z + 2) + \dfrac{3^2}{5^2}(z + 2)^2 + \dfrac{3^3}{5^3}(z + 2)^3 + \ldots\right]$, $R = \dfrac{5}{3}$. 57. $-iz + z^3 + iz^5 - z^7 - \ldots$, $R = 1$. 58. $\dfrac{1}{2}\left(\dfrac{z^2}{2!} + \dfrac{z^4}{4!} + \dfrac{z^6}{6!} + \ldots\right)$, $R = \infty$. 59. $z = 0$ (second order), $z_\pm = \pm 2i$ (simple). 60. $z_n = \pi n$ ($n = \pm1, \pm2, \ldots$) (simple). 61. $z = 0$ (third order), $z_n = \pi n$ ($n = \pm1, \pm2 \ldots$) (simple). 62. $z_n = (2n + 1)\pi i$ ($n = 0, \pm1, \pm2, \ldots$) (second order). 63. $|z| > e$

64. $2 < |z| < 4$. 65. $1 < |z| < 2$. 66. $0 < |z| < 1$. 67. $\frac{1}{z} - \frac{z}{3!} + \frac{z^3}{5!} - \frac{z^5}{7!} + \dots$

68. $\frac{1}{z^3} + \frac{1}{z^2} + \frac{1}{2!z} + \frac{1}{3!} + \frac{z}{4!} + \dots$. 69. $z^4 - \frac{z^2}{2!} + \frac{1}{4!} - \frac{1}{6!z^2} + \dots$. 70. $\frac{1}{z^2} - \frac{1}{2!z} + \frac{1}{3!} - \frac{z}{4!} + \dots$. 71. $\frac{1}{z} + \sum_{n=0}^{\infty} (-1)^n z^n$. 72. $\sum_{n=0}^{\infty} \frac{(-1)^n}{z^{n+2}}$. 73. $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{z^n} + \sum_{n=0}^{\infty} \frac{(-1)^n}{2^{n+1}} z^n$. 74. $\sum_{n=1}^{\infty} \frac{n4^{n-1}}{(z+2)^{n+1}}$. 75. $z = 0$ is a removable singularity. 76. $z = 0$ is an essential singularity. 77. $z = 0$ is a simple pole. 78. $z = 0$ is a pole of the fourth order, $z = -1$ is a simple pole. 79. $\operatorname{res} f(0) = -\frac{5}{2}$, $\operatorname{res} f(1) = e$. 80. $\operatorname{res} f(0) = 0$, 81. $\operatorname{res} f(0) = -\frac{1}{6}$. 82. $\operatorname{res} f(i) = -1$. 83. $\left(1 - \frac{2}{e}\right)\pi$. 84. $2\pi i$. 85. 0. 86. $2\pi i \frac{e^2}{3}$. 87. A removable singularity. 88. An essential singularity. 89. A removable singularity. 90. A pole of the third order. 91. $\frac{3\pi}{8}$. 92. $\frac{2\pi}{3}$. 93. $\frac{\pi}{2}$. 94. $\frac{\pi}{3e^3}$. 95. 0. 96. $\frac{\pi}{4e}$. 97. $\frac{b-a}{2}\pi$. 98. $\frac{2\pi}{1-p^2}$. 99. 0.

Chapter 27

Integral Transforms. Fourier Transforms

27.1 Fourier Integral

Integral transforms are a powerful tool in dealing with problems of mathematical physics.

Let a function $f(x)$ be defined on an interval (a, b) finite or infinite. The integral transform of $f(x)$ is the function

$$F(\omega) = \int_a^b K(x, \omega) f(x)\, dx, \tag{$*$}$$

where the function $K(x, \omega)$ is fixed for a given transformation; it is called the *kernel* of the transform (it is assumed that integral $(*)$ exists in the proper and improper senses).

Any function $f(x)$ which on an interval $[-l, l]$ meets the conditions for it to be expanded into a Fourier series can in the interval be represented by the trigonometric series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi}{l} x + b_n \sin \frac{n\pi}{l} x \right). \tag{27.1}$$

The coefficients a_n and b_n of (27.1) are defined by the Euler-Fourier formulas

$$\left.\begin{aligned} a_n &= \frac{1}{l} \int_{-l}^{l} f(\tau) \cos \frac{n\pi\tau}{l}\, d\tau \quad (n = 0, 1, 2, \ldots), \\ b_n &= \frac{1}{l} \int_{-l}^{l} f(\tau) \sin \frac{n\pi\tau}{l}\, d\tau \quad (n = 1, 2, \ldots) \end{aligned}\right\}. \tag{27.2}$$

The series on the right-hand side of (27.1) can be written in another form. So we will introduce into it from (27.2) the values of coefficients a_n and b_n, and bring under the sign of integrals $\cos (n\pi x/l)$ and $\sin (n\pi x/l)$ (which is possible because the integration variable here is τ) and use the formula

for the cosine of a difference. We will have

$$f(x) = \frac{1}{2l}\int_{-l}^{l} f(\tau)\,d\tau + \sum_{n=1}^{\infty}\frac{1}{l}\int_{-l}^{l} f(\tau)\cos\frac{n\pi(x-\tau)}{l}\,d\tau. \qquad (27.3)$$

If a function was originally defined on an interval of the number axis larger than the segment $[-l, l]$ (e.g., on the entire axis), then the expansion (27.3) will reproduce the values of the function only on the interval $[-l, l]$ and will continue it to the entire number axis as a periodic function with period $2l$ (Fig. 27.1). Therefore, if $f(x)$ is, generally speaking, a nonperiodic function, and it is defined on the entire number axis, in (27.3) we can try to

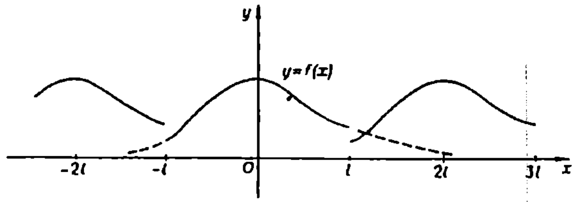


Fig. 27.1

go to the limit as $l \to +\infty$. It is quite natural to require that the following conditions be met:

(1) $f(x)$ meets the conditions for it to be expandable into a Fourier series on any finite segment on the x-axis;

(2) $f(x)$ is absolutely integrable on the entire number axis, i.e.,

$$\int_{-\infty}^{+\infty} |f(x)|\,dx = K < +\infty. \qquad (27.4)$$

If condition (27.4) is met, the first term on the right of (27.3) tends to zero as $l \to +\infty$. Indeed,

$$\left|\frac{1}{2l}\int_{-l}^{l} f(\tau)\,d\tau\right| \leqslant \frac{1}{2l}\int_{-l}^{l} |f(\tau)|\,d\tau \leqslant \frac{1}{2l}\int_{-\infty}^{+\infty} |f(\tau)|\,d\tau = \frac{K}{2l}\underset{l\to+\infty}{\longrightarrow} 0.$$

We will find the limit of the sum on the right of (27.3) as $l \to +\infty$. We put in (27.3)

$$\xi_1 = \frac{\pi}{l},\quad \xi_2 = \frac{2\pi}{l},\ \ldots,\ \xi_n = \frac{n\pi}{l},\ \ldots;$$

$\Delta\xi_n = \xi_{n+1} - \xi_n = \pi/l$, so that $1/l = \Delta\xi_n/\pi$.

The sum on the right of (27.3) will then become

$$\frac{1}{\pi}\sum_{n=1}^{\infty} \Delta\xi_n \int_{-l}^{l} f(\tau) \cos \xi_n(x-\tau)d\tau. \tag{27.5}$$

The integral converges absolutely, and so this sum for sufficiently large l will differ but slightly from the expression

$$\frac{1}{\pi}\sum_{n=1}^{\infty} \Delta\xi_n \int_{-\infty}^{+\infty} f(\tau) \cos \xi_n(x-\tau)d\tau,$$

which looks like an integral sum for a function of the variable ξ

$$\psi(\xi) = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(\tau) \cos \xi\,(x-\tau)d\tau$$

formed for the interval $(0, +\infty)$. Therefore, it would be only natural to expect that as $l \to +\infty$, when $\Delta\xi_n = \pi/l \to 0$, the sum (27.5) will become the integral

$$\int_{0}^{+\infty} \psi(\xi)d\xi = \frac{1}{\pi}\int_{0}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \cos \xi(x-\tau)d\tau.$$

On the other hand, as $l \to +\infty$ (x is fixed) it follows from (27.3) that

$$f(x) = \lim_{l\to+\infty}\sum_{n=1}^{\infty}\frac{1}{l}\int_{-l}^{l} f(\tau)\cos\frac{n\pi}{l}(x-\tau)d\tau, \tag{27.6}$$

and we obtain*)

$$f(x) = \frac{1}{\pi}\int_{0}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \cos \xi(x-\tau)d\tau. \tag{27.7}$$

A sufficient condition for (27.7) to be valid is given by the following.

Theorem 27.1. *If a function $f(x)$ is absolutely integrable on the interval $-\infty < x < +\infty$ and together with its derivative it has a finite number of discontinuities of the first kind on any finite interval $[a, b]$, then*

$$f(x) = \frac{1}{\pi}\int_{0}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \cos \xi(x-\tau)d\tau.$$

*) The reasoning is not rigorous.

We consider then that at any point x_0, which is a discontinuity of the first kind for $f(x)$, the integral on the right-hand side of (27.7) is equal to

$$\frac{1}{2}[f(x_0 - 0) + f(x_0 + 0)].$$

Formula (27.7) is known as the *Fourier integral formula*, and the integral on the right, as the *Fourier integral.*

If we make use of the formula for the cosine of a difference, formula (27.7) can be rewritten as

$$f(x) = \int_0^{+\infty} [a(\xi) \cos \xi x + b(\xi) \sin \xi x] d\xi, \tag{27.8}$$

where

$$a(\xi) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(\tau) \cos \xi\tau \, d\tau; \quad b(\xi) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(\tau) \sin \xi\tau \, d\tau \tag{27.9}$$

The functions $a(\xi)$, $b(\xi)$ are counterparts of the corresponding Fourier coefficients a_n and b_n of a function with period 2π, but these latter are defined for discrete values of n, whereas $a(\xi)$ and $b(\xi)$ are defined for continuous values of $\xi \in (-\infty, +\infty)$.

Complex form of the Fourier integral. Assuming $f(x)$ to be absolutely integrable on the entire x-axis, we will consider the integral

$$\int_{-\infty}^{+\infty} f(\tau) \sin \xi(x - \tau) d\tau, \quad -\infty < \xi < +\infty.$$

This integral converges uniformly for $-\infty < \xi < +\infty$, since

$$|f(\tau) \sin \xi(x - \tau)| \leqslant |f(\tau)|.$$

Therefore, it is a continuous and, obviously, odd function of ξ. But then

$$\lim_{l \to +\infty} \int_{-l}^{l} d\xi \int_{-\infty}^{+\infty} f(\tau) \sin \xi(x - \tau) d\tau = \int_{-\infty}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \sin \xi(x - \tau) d\tau = 0.$$

On the other hand, the integral

$$\int_{-\infty}^{+\infty} f(\tau) \cos \xi(x - \tau) d\tau, \quad -\infty < \xi < +\infty$$

is an even function of ξ, so that

$$\int_0^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \cos \xi(x - \tau) d\tau = \frac{1}{2} \int_{-\infty}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \cos \xi(x - \tau) d\tau.$$

Therefore, the Fourier integral formula can be written as

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \cos \xi(x - \tau) d\tau. \tag{27.10}$$

We multiply

$$0 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \sin \xi(x - \tau) d\tau$$

by the imaginary unit i and add the result to (27.10) to obtain

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau)[\cos \xi(x - \tau) + i \sin \xi(x - \tau)] d\tau.$$

Using the Euler formula $e^{i\varphi} = \cos \varphi + i \sin \varphi$, we will have

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) e^{i\xi(x - \tau)} d\tau = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\xi x} d\xi \int_{-\infty}^{+\infty} f(\tau) e^{-i\xi\tau} d\tau. \tag{27.11}$$

This is the *complex form of the Fourier integral.*

The external integration with respect to ξ is here understood in the sense of the Cauchy principal value

$$\int_{-\infty}^{+\infty} e^{i\xi x} \left(\int_{-\infty}^{+\infty} f(\tau) e^{-i\xi\tau} d\tau \right) d\xi = \lim_{N \to +\infty} \int_{-N}^{N} e^{i\xi x} \left(\int_{-\infty}^{+\infty} f(\tau) e^{-i\xi\tau} d\tau \right) d\xi.$$

27.2 Fourier Transform. Fourier Sine and Cosine Transforms

Let a function $f(x)$ be piecewise smooth on any finite segment of the x-axis and absolutely integrable on the axis.

Definition. A function

$$F(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx \tag{27.12}$$

is called the *Fourier transform* of $f(x)$ (spectral function). This is an integral transform of $f(x)$ on the interval $(-\infty, +\infty)$ with the kernel $K(x, \xi) = e^{-i\xi x}/\sqrt{2\pi}$.

Using the Fourier integral formula

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\xi x} d\xi \int_{-\infty}^{\infty} f(\tau) e^{-i\xi\tau} d\tau,$$

we will get

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(\xi) e^{i\xi x} d\xi. \tag{27.13}$$

This is the so-called *inverse Fourier transform* relating $F(\xi)$ to $f(x)$.

Sometimes the direct Fourier transform is given by

$$F(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{i\xi x} dx. \tag{27.12'}$$

Then, the inverse Fourier transform will be

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(\xi) e^{-i\xi x} d\xi. \tag{27.13'}$$

The Fourier transform $F(\xi)$ of $f(x)$ is also defined by

$$F(\xi) = \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx. \tag{27.12''}$$

Then in turn

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\xi) e^{i\xi x} d\xi. \tag{27.13''}$$

The position of the factor $1/2\pi$ is here sufficiently arbitrary: it can either enter into (27.12″) or (27.13″).

Examples. (1) Find the Fourier transform of the function $f(x) = e^{-\alpha x^2}$ $(\alpha > 0)$.

◄ By definition we have

$$F(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-i\xi x} dx. \tag{27.14}$$

We can here differentiate with respect to ξ under the integration sign (the resultant integral converges uniformly, when ξ belongs to any finite

segment)

$$F'(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-i\xi x}(-ix)\,dx.$$

Integrating by parts gives

$$F'(\xi) = \frac{1}{\sqrt{2\pi}} \frac{i}{2\alpha} \int_{-\infty}^{+\infty} e^{-i\xi x} d(e^{-\alpha x^2})$$

$$= \frac{1}{\sqrt{2\pi}} \left\{ \frac{i}{2\alpha} [e^{-i\xi x} e^{-\alpha x^2}]\Big|_{x=-\infty}^{x=+\infty} - \frac{i}{2\alpha} \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-i\xi x}(-i\xi)\,dx \right\}.$$

The first term vanishes and so we have

$$F'(\xi) = -\frac{\xi}{2\alpha} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{-i\xi x}\,dx = -\frac{\xi}{2\alpha} F(\xi),$$

hence

$$F(\xi) = Ce^{-\frac{\xi^2}{4\alpha}}, \tag{27.15}$$

where C is a constant of integration.

Putting in (27.15) $\xi = 0$ gives $C = F(0)$. By (27.14),

$$F(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\alpha x^2}\,dx. \tag{27.16}$$

It is well known that

$$\int_{-\infty}^{+\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

Therefore,

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2}\,dx = \sqrt{\frac{\pi}{\alpha}},$$

and hence, by (27.16), $F(0) = 1/\sqrt{2\alpha}$.

Thus

$$F(\xi) = \frac{1}{\sqrt{2\alpha}} e^{-\frac{\xi^2}{4\alpha}}.$$

Specifically, for $f(x) = e^{-x^2/2}$ $(\alpha = 1/2)$ we have

$$F(\xi) = e^{-\frac{\xi^2}{2}}. \blacktriangleright$$

(2) (Discharge of an electric capacitor through a resistor.) Consider the function

$$f(t) = \begin{cases} e^{-\alpha t}, & t > 0 \\ 0 & t < 0 \end{cases} \quad (\alpha > 0).$$

The spectral function $F(\xi)$ will be

$$F(\xi) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\alpha t} e^{-i\xi t} dt = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-(\alpha + i\xi)t} dt = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\alpha + i\xi} \cdot$$

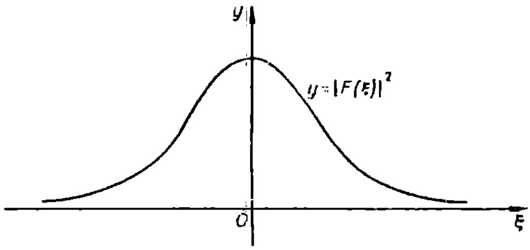


Fig. 27.2

Hence (Fig. 27.2)

$$|F(\xi)|^2 = \frac{1}{2\pi(\alpha^2 + \xi^2)} \cdot \blacktriangleright$$

The condition for $f(x)$ to be absolutely integrable on the interval $(-\infty, +\infty)$ is quite stringent. It excludes, for instance, such elementary functions as $f(x) = 1$, $f(x) = x^3$, $f(x) = \cos x$, $f(x) = e^x$, for which there exist no Fourier transforms (in classic form).

A Fourier image can only be found for functions that tend to zero as $|x| \to +\infty$ sufficiently fast (as in Examples (1) and (2)).

Fourier sine and cosine transforms. Using the formula for a cosine of a difference, we will rewrite the Fourier integral formula

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau) \cos \xi(x - \tau) d\tau$$

in the form

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} \cos \xi x \left(\int_{-\infty}^{+\infty} f(\tau) \cos \xi\tau \, d\tau \right) d\xi$$

$$+ \frac{1}{\pi} \int_0^{\infty} \sin \xi x \left(\int_{-\infty}^{+\infty} f(\tau) \sin \xi\tau \, d\tau \right) d\xi. \tag{27.17}$$

Let $f(x)$ be an even function. Then

$$\int_{-\infty}^{+\infty} f(\tau) \cos \xi\tau \, d\tau = 2 \int_0^{+\infty} f(\tau) \cos \xi\tau \, d\tau,$$

$$\int_{-\infty}^{+\infty} f(\tau) \sin \xi\tau \, d\tau = 0,$$

so that we have from (27.17)

$$f(x) = \frac{2}{\pi} \int_0^{+\infty} \cos \xi x \left(\int_0^{+\infty} f(\tau) \cos \xi\tau \, d\tau \right) d\xi. \tag{27.18}$$

If $f(x)$ is odd, we likewise obtain

$$f(x) = \frac{2}{\pi} \int_0^{+\infty} \sin \xi x \left(\int_0^{+\infty} f(\tau) \sin \xi\tau \, d\tau \right) d\xi. \tag{27.19}$$

If $f(x)$ is oniy defined on $[0, +\infty)$, then formula (27.18) extends $f(x)$ to the entire axis in an even manner, and (27.19) in an odd manner.

Definitions. The function

$$F_c(\xi) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} f(x) \cos \xi x \, dx \tag{27.20}$$

is called the *Fourier cosine transform* of a function $f(x)$.

It follows from (27.18) that

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} F_c(\xi) \cos \xi x \, d\xi. \tag{27.21}$$

This implies that $f(x)$ is in turn a cosine transform for $F_c(\xi)$. In other words, the functions f and F_c are mutual cosine transforms.

The function

$$F_s(\xi) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} f(x) \sin \xi x \, dx \tag{27.22}$$

is called the *Fourier sine transform* of $f(x)$.

By (27.19),

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} F_s(\xi) \sin \xi x \, d\xi, \tag{27.23}$$

i.e., f and F_s are mutual sine transforms.

Example. (Rectangular pulse.) Let $f(t)$ be an even function defined as follows:

$$f(t) = \begin{cases} 1 & \text{for } |t| < \theta \\ 0 & \text{for } |t| > \theta \end{cases} \quad (\theta > 0 = \text{const}).$$

Then

$$F_c(\xi) = \sqrt{\frac{2}{\pi}} \int_0^{\theta} \cos \xi t \, dt = \sqrt{\frac{2}{\pi}} \frac{\sin \xi\theta}{\xi}, \tag{27.24}$$

(see Fig. 27.3).

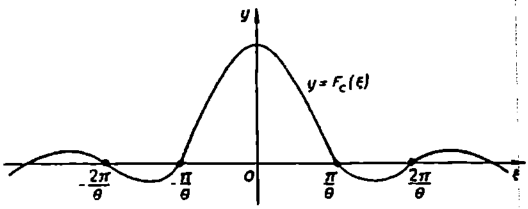


Fig. 27.3

We make use of these results to take the integral

$$\mathrm{Si}\,(\infty) = \int_0^{+\infty} \frac{\sin x}{x} \, dx.$$

By (27.21), we have

$$f(t) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} F_c(\xi) \cos \xi t \, d\xi = \frac{2}{\pi} \int_0^{+\infty} \frac{\sin \theta\xi}{\xi} \cos \xi t \, d\xi. \tag{27.24'}$$

At $t = 0$ the function $f(t)$ is continuous and equal to unity. Therefore, from (27.24′),

$$1 = \frac{2}{\pi}\int_0^{+\infty} \frac{\sin\theta\xi}{\xi}\,d\xi = \frac{2}{\pi}\int_0^{+\infty} \frac{\sin\theta\xi}{\theta\xi}\,d(\theta\xi) = \frac{2}{\pi}\int_0^{+\infty} \frac{\sin x}{x}\,dx,$$

so that

$$\int_0^{+\infty} \frac{\sin x}{x}\,dx = \frac{\pi}{2}. \blacktriangleright$$

Amplitude and phase spectra of the Fourier integral. Suppose that a function $f(x)$ with period 2π is expandable into a Fourier series

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

We can represent this in the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} c_n \cos(nx - \varphi_n),$$

where $c_n = \sqrt{a_n^2 + b_n^2}$ is the amplitude of oscillations with frequency n, and φ_n is the phase. We thus come to the concept of the amplitude and phase spectra of a periodic function.

Under appropriate conditions a nonperiodic function $f(x)$ defined on $(-\infty, +\infty)$ can be represented by a Fourier integral

$$f(x) = \frac{1}{\pi}\int_0^{+\infty} d\xi \int_{-\infty}^{+\infty} f(\tau)\cos\xi(x - \tau)\,d\tau,$$

which is an expansion of the function in all the frequencies $0 < \xi < +\infty$ (expansion in a continuous frequency spectrum).

Definition. The *spectral function* or the *spectral density* of a Fourier integral is the expression

$$S(\xi) = \int_{-\infty}^{+\infty} f(x)e^{-i\xi x}\,dx$$

(the direct Fourier transform of $f(x)$).

The function

$$A(\xi) = |S(\xi)| = \sqrt{\left(\int_{-\infty}^{+\infty} f(x)\cos \xi x\, dx\right)^2 + \left(\int_{-\infty}^{+\infty} f(x)\sin \xi x\, dx\right)^2}$$

is called the *amplitude spectrum*, and the function

$$\Phi(\xi) = -\arg S(\xi)$$

is called the *phase spectrum* of $f(x)$.

The amplitude spectrum $A(\xi)$ is a measure of the contribution of frequency ξ to $f(x)$.

Example. Find the amplitude and phase spectra of the function $f(x) = e^{-|x|}$.

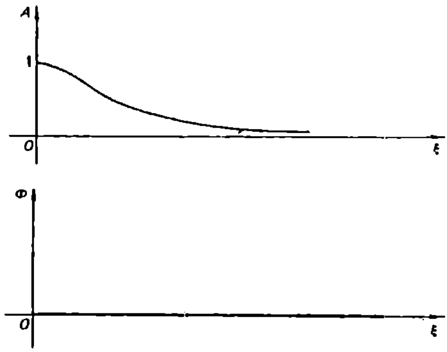


Fig. 27.4

◄ We find the spectral function

$$S(\xi) = \int_{-\infty}^{+\infty} f(x)e^{-i\xi x}dx = \int_{-\infty}^{0} e^{x}e^{-i\xi x}dx + \int_{0}^{+\infty} e^{-x}e^{-i\xi x}dx$$

$$= \frac{e^{x(1-i\xi)}}{1-i\xi}\Bigg|_{x=-\infty}^{x=0} - \frac{e^{-x(1+i\xi)}}{1+i\xi}\Bigg|_{x=0}^{x=+\infty} = \frac{1}{1+\xi^2}.$$

Hence $A(\xi) = |S(\xi)| = 1/(1+\xi^2)$, $\Phi(\xi) = -\arg S(\xi) = 0$. The graphs of the functions are given in Fig. 27.4. ►

27.3 Properties of the Fourier Transform

(1) *Linearity.* If $F(\xi)$ and $G(\xi)$ are the Fourier transforms of $f(x)$ and $g(x)$, respectively, then at any constants α and β, the Fourier transform of $\alpha f(x) + \beta g(x)$ will be the function $\alpha F(\xi) + \beta G(\xi)$.

Indeed, using the linearity property of the integral we have

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\alpha f(x)+\beta g(x))e^{-i\xi x}dx=\alpha\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}f(x)e^{-i\xi x}dx$$

$$+\beta\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}g(x)e^{-i\xi x}dx=\alpha F(\xi)+\beta G(\xi).$$

Accordingly, the Fourier transform is a linear operator. Denoting it by $\mathscr{F}$ we write

$$f(x)\overset{\mathscr{F}}{\rightarrow}F(\xi)\quad\text{or}\quad\mathscr{F}[f(x)]=F(\xi).$$

(2) If $F(\xi)$ is the Fourier transform of a function $f(x)$ that is absolutely integrable on the interval $(-\infty, +\infty)$, then $F(\xi)$ is bounded for all $\xi\in(-\infty, +\infty)$.

◄ Let $f(x)$ be absolutely integrable for $-\infty<x<+\infty$:

$$\int_{-\infty}^{+\infty}|f(x)|dx=K<+\infty,$$

and let

$$F(\xi)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}f(x)e^{-i\xi x}dx$$

be the Fourier transform of $f(x)$. Then

$$|F(\xi)|=\frac{1}{\sqrt{2\pi}}\left|\int_{-\infty}^{+\infty}f(x)e^{-i\xi x}dx\right|\leqslant\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}|f(x)|dx=\frac{K}{\sqrt{2\pi}},$$

which proves the statement. ►

Problems. (1). Let $f(x)$ be a function that can be Fourier transformed, and h be a real number. The function $f_h(x)=f(x-h)$ is called the shift of $f(x)$. Using the definition of the Fourier transform, show that

$$\mathscr{F}[f_h]=e^{-i\xi h}\mathscr{F}[f].$$

(2) Let $f(x)$ have a Fourier transform $F(\xi)$, h being a real number. Show that

$$\mathscr{F}[e^{ihx}f(x)]=F(\xi-h).$$

(3) *Fourier transform and the operation of differentiation.* Suppose that an absolutely integrable function $f(x)$ has a derivative $f'(x)$ that is also

absolutely integrable for $-\infty < x < +\infty$, so that $f(x)$ tends to zero as $|x| \to +\infty$.

The condition $f(x) \underset{|x| \to +\infty}{\longrightarrow} 0$ is fairly natural, since, according to N. Wiener, the usual theory of Fourier integrals deals with processes that in some sense or other have a beginning and an end, but are not unboundedly extendable with about the same intensity.

Assuming $f'(x)$ to be a smooth function, we write

$$\mathcal{F}[f'] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f'(x) e^{-i\xi x} dx.$$

Integrating by parts gives

$$\mathcal{F}[f'] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f'(x) e^{-i\xi x} dx$$

$$= \frac{1}{\sqrt{2\pi}} \left[f(x) e^{-i\xi x} \Big|_{x=-\infty}^{x=+\infty} + i\xi \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx \right].$$

The first term vanishes (since $f(x) \to 0$ as $|x| \to +\infty$), and we arrive at

$$\mathcal{F}[f'] = i\xi \mathcal{F}[f]. \tag{27.25}$$

And so *to differentiate $f(x)$ means to multiply its Fourier image $\mathcal{F}[f]$ by $i\xi$.*

If $f(x)$ has smooth absolutely integrable derivatives up to order m, and they all, just like $f(x)$ itself, tend to zero as $|x| \to +\infty$, then integrating by parts the required number of times, we will obtain

$$\mathcal{F}[f^{(k)}(x)] = (i\xi)^k \mathcal{F}[f(x)] \quad (k = 0, 1, \ldots, m). \tag{27.26}$$

The Fourier transform is very useful precisely because it replaces the operation of differentiation by the operation of multiplication by the quantity $i\xi$, thereby simplifying the task of integrating some types of differential equations.

Since the Fourier transform $\mathcal{F}[f^{(k)}(x)]$ of an absolutely integrable function $f^{(k)}(x)$ is a bounded function of ξ (Property (2)), from (27.26) we find

$$|\mathcal{F}[f]| = \frac{|\mathcal{F}[f^{(k)}]|}{|\xi|^k} \leqslant \frac{C}{|\xi|^k} \quad (C > 0 = \text{const}).$$

It follows from this estimate that *the more absolutely integrable derivatives has a function $f(x)$ the faster its Fourier transform tends to zero as $|\xi| = +\infty$.*

(4) *Correlation between the rate of decreasing of $f(x)$ as $|x| \to +\infty$ and the smoothness of its Fourier transform.*

Suppose that not only $f(x)$ but also the product $x \cdot f(x)$ are absolutely integrable functions on the entire x-axis.

Then the Fourier transform for $f(x)$

$$F(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx$$

will be differentiable. Indeed, formal differentiation with respect to the parameter ξ of the integrand leads to the integral

$$-\frac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x f(x) e^{-i\xi x} dx,$$

which is absolutely and uniformly convergent in ξ. Correspondingly, the differentiation is legitimate and

$$F'(\xi) = -\frac{i}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x f(x) e^{-i\xi x} dx = -i\mathcal{F}[xf(x)].$$

Thus

$$i \frac{d}{d\xi} F(\xi) = \mathcal{F}[xf(x)],$$

i.e., the operation of multiplication of $f(x)$ by the argument x after the Fourier transformation changes to the operation $i \dfrac{d}{d\xi}$.

If along with $f(x)$, the functions $xf(x), \ldots, x^m f(x)$ are absolutely integrable on the entire x-axis, then the differentiation process can be carried on. We will find that the function $F(\xi) = \mathcal{F}[f(x)]$ yields derivatives up to order m, and

$$i^k \frac{d^k}{d\xi^k} \mathcal{F}[f] = \mathcal{F}[(x^k f(x)] \quad (k = 0, 1, \ldots, m).$$

Thus, *the faster a function $f(x)$ decreases as $|x| \to +\infty$, the more smooth the function $F(\xi) = \mathcal{F}[f(x)]$ is.*

(5) **Theorem 27.2 (convolution theorem).** Let $F_1(\xi)$ and $F_2(\xi)$ be Fourier transforms of the functions $f_1(x)$ and $f_2(x)$, respectively. Then

$$F_1(\xi) F_2(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f_1(x) e^{-i\xi x} dx \, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f_2(y) e^{-i\xi y} dy$$

$$= \frac{1}{(\sqrt{2\pi})^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1(x) f_2(y) e^{-i\xi(x+y)} dx\, dy,$$

where the double integral on the right-hand side converges absolutely.

We put $x + y = \tau$, and so $y = \tau - x$. Then

$$F_1(\xi)F_2(\xi) = \frac{1}{(\sqrt{2\pi})^2} \int_{-\infty}^{+\infty} f_1(x) \left\{ \int_{-\infty}^{+\infty} f_2(\tau - x) e^{-i\xi\tau} d\tau \right\} dx.$$

If we change the order of integration, we will get

$$F_1(\xi)F_2(\xi) = \frac{1}{(\sqrt{2\pi})^2} \int_{-\infty}^{+\infty} e^{-i\xi\tau} \left\{ \int_{-\infty}^{+\infty} f_1(x) f_2(\tau - x) dx \right\} d\tau. \qquad (27.27)$$

The function

$$\varphi(\tau) = \int_{-\infty}^{+\infty} f_1(x) f_2(\tau - x) dx, \quad -\infty < \tau < +\infty$$

is called the *convolution* of $f_1(x)$ and $f_2(x)$, and is denoted by $(f_1 * f_2)(\tau)$.

Formula (27.27) can then be written as

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(\tau) e^{-i\xi\tau} d\tau = \sqrt{2\pi} F_1(\xi) F_2(\xi).$$

It is seen that *the Fourier transform of the convolution of $f_1(x)$ and $f_2(x)$ is equal to the product of the Fourier transforms of the convoluted functions and $\sqrt{2\pi}$*

$$\mathcal{F}[f_1 * f_2] = \sqrt{2\pi} \mathcal{F}[f_1] \cdot \mathcal{F}[f_2].$$

Remark. We can readily establish the following properties of the convolution:

(1) *Linearity*: $f * (\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 (f * f_1) + \alpha_2 (f * f_2)$, where α_1 and α_2 are constants;

(2) *Commutativity*: $f_1 * f_2 = f_2 * f_1$.

27.4 Applications

(1) Let $P\left(\frac{d}{dx}\right)$ be a linear differential operator of order m with the constant coefficients

$$P\left(\frac{d}{dx}\right) \equiv a_0 \frac{d^m}{dx^m} + a_1 \frac{d^{m-1}}{dx^{m-1}} + \ldots + a_{m-1} \frac{d}{dx} + a_m,$$

where $a_0, a_1, \ldots, a_m$ are constants. Using the formula for the Fourier transform of the derivatives of $y(x)$, we find

$$\mathcal{F}\left[P\left(\frac{d}{dx}\right)y\right] = P(i\xi)\mathcal{F}[y].$$

Consider the differential equation

$$P\left(\frac{d}{dx}\right)y = f(x), \tag{27.28}$$

where $P\left(\frac{d}{dx}\right)$ is a differential operator introduced above.

We assume that the solution $y(x)$ we seek has the Fourier transform $\tilde{y}(\xi)$, and the function $f(x)$ has a transform $\tilde{f}(\xi)$. Applying the Fourier transform (27.28), we obtain instead of the differential equation an algebraic equation in $\tilde{y}(\xi)$ defined on the ξ-axis:

$$P(i\xi)\tilde{y} = \tilde{f}(\xi),$$

hence

$$\tilde{y}(\xi) = \frac{\tilde{f}(\xi)}{P(i\xi)},$$

so that formally we get

$$y(x) = \mathcal{F}^{-1}\left\{\frac{\tilde{f}(\xi)}{P(i\xi)}\right\},$$

where $\mathcal{F}^{-1}$ stands for the inverse Fourier transform.

The main limitation of the method is associated with the following fact. The solution of an ordinary differential equation with constant coefficients contains functions of the type $e^{\lambda x}$, $e^{\alpha x}\cos\beta x$, $e^{\alpha x}\sin\beta x$. They are not absolutely integrable on the axis $-\infty < x < +\infty$ and the Fourier transform for them is not defined, so that, strictly speaking, we may not apply the method. This limitation can be circumvented by introducing the so-called generalized functions. But in some cases we can still apply the Fourier transform in its classic form.

Example. Find the solution $u = u(x, t)$ of the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad -\infty < x < +\infty, \tag{27.29}$$

where $a = \text{const}$, under the initial conditions

$$u|_{t=0} = \varphi(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = 0, \quad -\infty < x < +\infty. \tag{27.30}$$

This is the problem of free oscillations of the uniform infinite string, when we know the initial displacement $\varphi(x)$ of points of the string but not the initial velocities.

Since the spatial variable x varies from $-\infty$ to $+\infty$, we will subject the equation and the initial conditions to Fourier transformation in x.

We will assume that:

(1) functions $u(x, t)$ and $\varphi(x)$ are sufficiently smooth and tend to zero as $|x| \to +\infty$ and $\forall t \geqslant 0$ so fast that there exists the Fourier transform

$$v(\xi, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u(x, t) e^{-i\xi x} dx, \tag{27.31}$$

$$\tilde{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x) e^{-i\xi x} dx. \tag{27.32}$$

(2) differentiation is possible, so that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\partial^2 u}{\partial t^2} e^{-i\xi x} dx = \frac{\partial^2}{\partial t^2} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u(x, t) e^{-i\xi x} dx \right) = \frac{d^2 v(\xi, t)}{dt^2}, \tag{27.33}$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\partial^2 u}{\partial x^2} e^{-i\xi x} dx = (i\xi)^2 v(\xi, t) = -\xi^2 v(\xi, t). \tag{27.34}$$

Multiplying both sides of (27.29) by $e^{-i\xi x}/\sqrt{2\pi}$ and integrating with respect to x from $-\infty$ to $+\infty$, we will obtain

$$\frac{d^2 v}{dt^2} + a^2 \xi^2 v = 0, \tag{27.35}$$

and from the initial conditions (27.30) we will find

$$v|_{t=0} = \tilde{\varphi}(\xi), \tag{27.36}$$

$$\left. \frac{dv}{dt} \right|_{t=0} = 0. \tag{27.37}$$

If thus we apply to the problem (27.29-30) the Fourier transformation, we will come to the Cauchy problem (27.35-37) for an ordinary differential equation, where ξ is a parameter.

The solution of (27.35) is the function

$$v(\xi, t) = C_1(\xi) \cos a\xi t + C_2(\xi) \sin a\xi t.$$

We find from the conditions (27.36) and (27.37) that $C_1(\xi) = \tilde{\varphi}(\xi)$, $C_2(\xi) = 0$, so that $v(\xi, t) = \tilde{\varphi}(\xi) \cos a\xi t$. Applying inverse Fourier transformation, we will get

$$u(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} v(\xi, t) e^{i\xi x} d\xi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{\varphi}(\xi) \cos a\xi t \, e^{i\xi x} d\xi$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{\varphi}(\xi) \left[\frac{e^{i\xi(x + at)} + e^{i\xi(x - at)}}{2} \right] d\xi$$

$$= \frac{\varphi(x + at) + \varphi(x - at)}{2} .$$

This is a special case of the D'Alembert solution of the problem (27.29-30) (see Chap. 30).

(2) Fourier transforms can be used to solve some integral equations, i.e., equations in which the unknown function appears under the integration sign.

Consider, for example, the equation

$$\int_{-\infty}^{+\infty} \varphi(x) e^{-i\xi x} dx = 2\pi e^{-|\xi|}, \tag{27.38}$$

where $\varphi(x)$ is the desired function. Having written (27.38) in the form

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x) e^{-i\xi x} dx = \sqrt{2\pi} e^{-|\xi|}, \tag{27.39}$$

we notice that the left-hand side of this relation can be regarded as the Fourier transform of the function $\varphi(x)$, so that (27.39) is equivalent to

$$\mathscr{F}[\varphi] = \sqrt{2\pi} e^{-|\xi|}.$$

Then, by the inversion formula

$$\varphi(x) = \int_{-\infty}^{+\infty} e^{-|\xi|} e^{i\xi x} d\xi = \int_{-\infty}^{0} e^{\xi(1 + ix)} d\xi + \int_{0}^{+\infty} e^{-\xi(1 - ix)} d\xi$$

$$= \frac{e^{\xi(1 + ix)}}{1 + ix} \Bigg|_{\xi = -\infty}^{\xi = 0} - \frac{e^{-\xi(1 - ix)}}{1 - ix} \Bigg|_{\xi = 0}^{\xi = +\infty} = \frac{2}{1 + x^2} .$$

The function $\varphi(x) = 2/(1 + x^2)$ is a solution of (27.38).

27.5 Multiple Fourier Transforms

Let $x = (x_1, x_2, \ldots, x_n) \in R^n$. We define the Fourier transform of the absolutely integrable function $f(x_1, x_2, \ldots, x_n)$ by

$$F(\xi_1, \xi_2, \ldots, \xi_n)$$

$$= \frac{1}{(\sqrt{2\pi})^n} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, x_2, \ldots, x_n) e^{-i(\xi_1 x_1 + \xi_2 x_2 + \xi_n x_n)} dx_1 dx_2 \ldots dx_n$$

or, for short,

$$F(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{R^n} f(x) e^{-i(\xi, x)} dx,$$

where $(\xi, x) = \xi_1 x_1 + \xi_2 x_2 + \ldots + \xi_n x_n$, $dx = dx_1 dx_2 \ldots dx_n$. Integration here is over the whole of R^n.

The properties of the n-dimensional transform are similar to the appropriate properties of the Fourier transform of a function of one variable.

In the special case of $f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2) \ldots f(x_n)$ we have

$$\mathscr{F}[f_1, f_2 \ldots f_n] = \mathscr{F}[f_1]\mathscr{F}[f_2] \ldots \mathscr{F}[f_n].$$

Table 27.1 Fourier transforms

$f(x)$	$F(\xi) = \mathscr{F}[f] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\xi x} dx$
$f(ax)$, $a > 0$	$\frac{1}{a} F\left(\frac{\xi}{a}\right)$
$f(x - a)$	$e^{-i\xi a} F(\xi)$
$e^{-a^2 x^2}$	$\frac{1}{a\sqrt{2}} e^{-\xi^2/4a^2}$
$e^{-a\lvert x\rvert}$, $a > 0$	$\sqrt{\frac{2}{\pi}} \frac{a}{a^2 + \xi^2}$

Table 27.1 (concluded)

$f(x)$	$F(\xi) = \mathcal{F}[f] = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} f(x) e^{-i\xi x}\, dx$
$\begin{cases} 1, & \lvert x\rvert < a \\ 0, & \lvert x\rvert > a \end{cases}$	$\sqrt{\dfrac{2}{\pi}}\, \dfrac{\sin a\xi}{\xi}$
$\dfrac{1}{1+x^2}$	$\sqrt{\dfrac{\pi}{2}}\, e^{-\lvert\xi\rvert}$
$\dfrac{a}{x^2+a^2}$	$\sqrt{\dfrac{\pi}{2}}\, e^{-a\lvert\xi\rvert}$
$x e^{-a\lvert x\rvert}, \quad a > 0$	$-2\sqrt{\dfrac{2}{\pi}}\, \dfrac{ia\xi}{(\xi^2+a^2)^2}$
$\begin{cases} \cos ax, & \lvert x\rvert < \dfrac{\pi}{2a} \\ 0, & \lvert x\rvert > \dfrac{\pi}{2a} \end{cases}$	$\sqrt{\dfrac{2}{\pi}}\, \dfrac{a}{a^2-\xi^2} \cos \dfrac{\pi\xi}{2a}$
$\begin{cases} 1-\lvert x\rvert, & \lvert x\rvert < 1 \\ 0, & \lvert x\rvert > 1 \end{cases}$	$2\sqrt{\dfrac{2}{\pi}} \left(\dfrac{\sin \xi/2}{\xi}\right)^2$

Exercises

Represent by the Fourier integral the following functions:

1. $f(x) = \begin{cases} 1, & \lvert x\rvert < 1, \\ \dfrac{1}{2}, & x = \pm 1, \\ 0, & \lvert x\rvert > 1. \end{cases}$ 2. $f(x) = \begin{cases} 1, & a < x < b, \\ 0, & x < a, \quad x > b. \end{cases}$

3. $f(x) = e^{-\lvert x\rvert}$. 4. $f(x) = \dfrac{x}{x^2+a^2}$, $a > 0$. 5. $f(x) = \begin{cases} \sin x, & \lvert x\rvert \leqslant \pi, \\ 0, & \lvert x\rvert > \pi. \end{cases}$

6. Solve the integral equation: $\displaystyle\int\limits_0^{+\infty} f(t) \cos \xi t\, dt = \frac{1}{1+\xi^2}$.

Answers

1. $f(x) = \frac{2}{\pi}\int_0^{+\infty} \frac{\sin\xi}{\xi}\cos\xi x\, d\xi$, $\int_0^{+\infty} \frac{\sin t}{t}\, dt = \frac{\pi}{2}$.

2. $f(x) = \frac{1}{\pi}\int_0^{+\infty} \frac{\sin\xi(x-a) - \sin\xi(x-b)}{\xi}\, d\xi$. 3. $f(x) = \frac{2}{\pi}\int_0^{+\infty} \frac{\cos\xi x}{1+\xi^2}\, d\xi$.

4. $f(x) = \int_0^{+\infty} e^{-a\xi}\sin x\xi\, d\xi$. 5. $f(x) = \frac{2}{\pi}\int_0^{+\infty} \frac{\sin\xi\pi}{1-\xi^2}\sin x\xi\, d\xi$. 6. $f(x) = e^{-x}$, $x \geqslant 0$.

Chapter 28

Laplace Transform

28.1 Basic Definitions

In Chapter 27 we were concerned with the integral Fourier transform

$$F(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-i\xi t} dt$$

with the kernel $K(t, \xi) = e^{-i\xi t}/\sqrt{2\pi}$.

The Fourier transform is inconvenient in that it is associated with the condition for the function

$$\int_{-\infty}^{+\infty} |f(t)| dt = A < +\infty \qquad (*)$$

to be absolutely integrable on the entire t-axis. The *Laplace transform* enables us to lift this constraint.

Definition. We define the *inverse transform* (the *original function*) to be any complex-valued function $f(t)$ of a real argument t obeying the following conditions:

(1) $f(t)$ is continuous on the entire t-axis, except for individual points at which $f(t)$ has discontinuities of the first kind, and in each finite interval of t there can only be a finite number of such points;

(2) $f(t)$ is equal to zero for negative t, i.e., $f(t) = 0$ for $t < 0$;

(3) as t increases, the modulus of $f(t)$ increases not faster than an exponential function, i.e., there exist numbers $M > 0$ and s, such that for all t

$$|f(t)| \leqslant Me^{st}. \qquad (28.1)$$

It is quite obvious that if (28.1) holds at some $s = s_1$, it will hold for any $s_2 > s_1$.

The greatest lower bound s_0 of all numbers s obeying (28.1) is called the *infimum* of $f(t)$, i.e., $s_0 = \inf s$. In the general case the inequality

$$|f(t)| \leqslant Me^{s_0 t}$$

does not hold, but does the inequality

$$|f(t)| \leqslant Me^{(s_0+\varepsilon)t},$$

where $\varepsilon > 0$ is any number. So, the function $f(t) = t$, $t \geqslant 0$, has the infimum $s_0 = 0$. The function does not obey the inequality $|t| \leqslant M$ for all $t \geqslant 0$, but instead we have $|t| \leqslant Me^{\varepsilon t}$ for all $\varepsilon > 0$ and $t \geqslant 0$.

Condition (28.1) is far less stringent than (*).

For example, the function

$$f(t) = \begin{cases} 0 & \text{for} \quad t < 0, \\ e^t & \text{for} \quad t \geqslant 0 \end{cases}$$

does not satisfy the condition (*), but condition (28.1) is met for any $s \geqslant 1$ and $M \geqslant 1$, and so $s_0 = 1$. Therefore, $f(t)$ is an inverse transform.

On the other hand, the function

$$f(t) = \begin{cases} 0 & \text{for} \quad t < 0, \\ e^{t^2} & \text{for} \quad t \geqslant 0 \end{cases}$$

is no inverse transform: it grows infinitely, i.e., $s_0 = +\infty$.

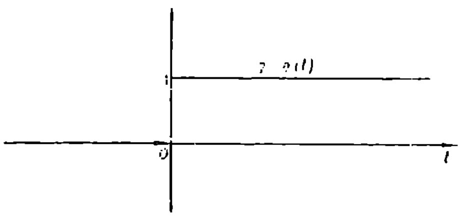


Fig. 28.1

The simplest inverse transform is the so-called *unit* (or *identity*) *function* $\eta(t)$ (Fig. 28.1):

$$\eta(t) = \begin{cases} 0 & \text{for} \quad t < 0, \\ 1 & \text{for} \quad t \geqslant 0. \end{cases}$$

If some function $\varphi(t)$ meets the conditions (1) and (3) of the definition but not (2), then the function $f(t) = \varphi(t)\eta(t)$ will already be an inverse transform. To simplify the notation we will as a rule omit the factor $\eta(t)$ on the assumption that all the functions that we will consider are equal to zero for negative t, since if we deal with some function $f(t)$, e.g., sin t,

$\cos t$, e^t, etc., then we always mean the following functions (Fig. 28.2):

$$f_1(t) = \begin{cases} 0 & \text{for } t < 0, \\ \sin t & \text{for } t \geqslant 0, \end{cases} \quad f_2(t) = \begin{cases} 0 & \text{for } t < 0, \\ \cos t & \text{for } t \geqslant 0, \end{cases}$$

$$f_3(t) = \begin{cases} 0 & \text{for } t < 0, \\ e^t & \text{for } t \geqslant 0. \end{cases}$$

Definition. Let $f(t)$ be an original function. The *Laplace transform* of $f(t)$ is a function $F(p)$ of the complex variable $p = s + i\sigma$ defined by

$$F(p) = \int_0^{+\infty} f(t)e^{-pt}dt, \tag{28.2}$$

where the integration is along the positive t-axis the transform kernel being $K(t, p) = e^{-pt}$.

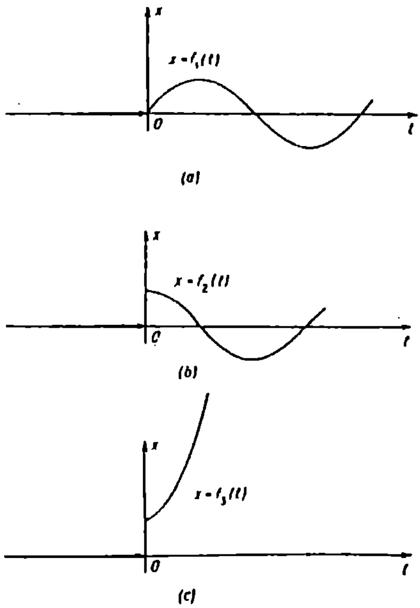


Fig. 28.2

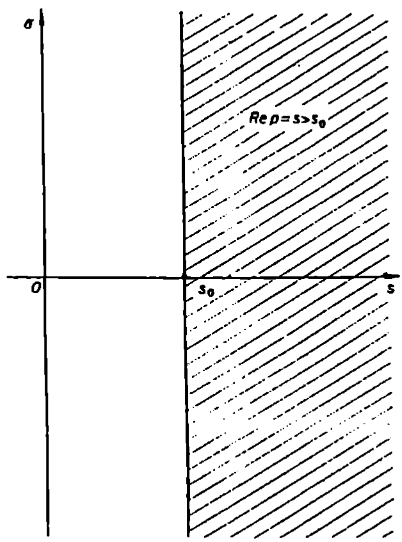


Fig. 28.3

We will use the following notation to indicate that a function $f(t)$ has a transform $F(p)$:

$$f(t) \risingdotseq F(p), \quad F(p) \fallingdotseq f(t) \quad \text{or} \quad F(p) = \mathscr{L}(f(t)).$$

Example. Find the Laplace transform of the unit function $\eta(t)$.

◄ The function $\eta(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t \geqslant 0 \end{cases}$ is an inverse transform with the infimum $s_0 = 0$. By (28.2), the transform of $\eta(t)$ will be

$$F(p) = \int_0^{+\infty} 1 \cdot e^{-pt} dt.$$

If $p = s + i\sigma$, then for $s > 0$ the integral on the right-hand side of this relation will converge, and so we will get

$$F(p) = \int_0^{\infty} 1 \cdot e^{-pt} dt = -\frac{e^{-pt}}{p}\Bigg|_{t=0}^{t=+\infty} = \frac{1}{p} \quad (\operatorname{Re} p = s > 0),$$

so that $\eta(t)$ will have as its transform the function $1/p$. We agreed to write that $\eta(t) = 1$ and then the result obtained will be

$$1 \risingdotseq \frac{1}{p}. \text{ ►}$$

Theorem 28.1. *The Laplace transform $F(p)$ of any inverse transform $f(t)$ with the infimum s_0 is defined in the half-plane* $\operatorname{Re} p = s > s_0$ (Fig. 28.3) *and is there an analytic function.*

◄ Let

$$|f(t)| \leqslant Me^{s_0 t}. \tag{28.3}$$

To prove that $F(p)$ exists in the indicated half-plane it is sufficient to establish that the improper integral (28.2) converges absolutely for $s > s_0$.

Using (28.3) gives

$$\left| \int_0^{+\infty} f(t) e^{-pt} dt \right| \leqslant \int_0^{+\infty} |f(t)| |e^{-(s+i\sigma)t}| dt$$

$$\leqslant M \int_0^{+\infty} e^{-(s-s_0)t} dt = \frac{M}{s - s_0},$$

which proves the absolute convergence of (28.2). At the same time this is an estimate of the Laplace transform $F(p)$ in the half-plane $\operatorname{Re} p = s > s_0$, which is called *the half-plane of convergence*

$$|F(p)| = \left| \int_0^{+\infty} f(t) e^{-pt} dt \right| \leqslant \frac{M}{s - s_0}. \tag{28.4}$$

Differentiating exprcsion (28.2) formally under the integration sign with respect to p, we find

$$F'(p) = - \int_0^{+\infty} t f(t) e^{-pt} dt. \tag{28.5}$$

The existence of integral (28.5) is established in the same way as of integral (28.2). So, if we integrate by parts, we obtain the estimate for $F'(p)$

$$|F'(p)| \leqslant M \int_0^{+\infty} t e^{-(s-s_0)t} dt$$

$$= M \left[\frac{t e^{-(s-s_0)t}}{-(s-s_0)} \Big|_{t=0}^{t=+\infty} + \frac{1}{s-s_0} \int_0^{+\infty} e^{-(s-s_0)t} dt \right] = \frac{M}{(s-s_0)^2}.$$

We infer from this that (28.5) converges absolutely. (The first term in the brackets has a limit equal to zero when $t \to +\infty$.)

In any half-plane $\operatorname{Re} p \geqslant s_1 > s_0$ the integral (28.5) converges uniformly in p, since it is dominated by the convergent integral

$$\left| \int_0^{+\infty} t f(t) e^{-pt} dt \right| \leqslant M \int_0^{+\infty} t e^{-(s_1-s_0)t} dt = \frac{M}{(s_1-s_0)^2},$$

which does not depend on p. Accordingly, it is legitimate to differentiate with respect to p, and so (28.5) holds.

Since the derivative $F'(p)$ does exist, the Laplace transform $F(p)$ everywhere in the half-plane $\operatorname{Re} p = s > s_0$ is an analytic function. ►

The inequality (28.4) has the following

Corollary. If a point p tends to infinity so that $\operatorname{Re} p = s$ grows unboundedly, then $F(p) \to 0$, i.e., $\lim\limits_{s \to +\infty} F(p) = 0$.

By way of example we find the transform of the function $f(t) = e^{at}$, where $a = \alpha + i\beta$ is any complex number. The infimum of $f(t)$ is equal to α, i.e., $s_0 = \alpha$.

If we assume that $\operatorname{Re} p = s > \alpha$, we will get

$$\int_0^{+\infty} e^{at} e^{-pt} dt = \int_0^{+\infty} e^{-(p-a)t} dt = \frac{e^{-(p-a)t}}{-(p-a)} \Big|_{t=0}^{t=+\infty} = \frac{1}{p-a}$$

($\operatorname{Re} p > \alpha$). Thus

$$e^{at} \risingdotseq \frac{1}{p-a}.$$

At $a = 0$ we again obtain the formula $1 \risingdotseq 1/p$.

Notice that the image of $e^{\alpha t}$ is an analytic function of p not only in the half-plane $\mathrm{Re}\, p > \alpha$, but also at all points p, except for the point $p = \alpha$, where the transform has a simple pole. Later in the chapter we will often encounter a situation where $F(p)$ is an analytic function in the entire plane of the complex variable p, except at some isolated singularities. There is no conflict with Theorem 28.1 here. The theorem only states that in the half-plane $\mathrm{Re}\, p > s_0$ the function $F(p)$ will no longer have singularities: they are all lying either to the left of the line $\mathrm{Re}\, p = s_0$ or on the line itself.

Remark. Operational calculus sometimes uses the Heaviside transform of a function $f(t)$, given by

$$F^*(p) = p \int_0^{+\infty} f(t) e^{-pt} dt.$$

Therefore, the Heaviside transform differs from the Laplace transform by factor p.

28.2 Properties of Laplace Transform

We define by $f(t)$, $\varphi(t)$, ... inverse transforms, and by $F(p)$, $\Phi(p)$, ... their Laplace transforms

$$F(p) = \int_0^{+\infty} f(t) e^{-pt} dt, \quad \Phi(p) = \int_0^{+\infty} \varphi(t) e^{-pt} dt, \tag{28.6}$$

Above all, it follows from the definition of the transform that if $f(t) \equiv 0$ for all t, then $F(p) \equiv 0$.

Theorem 28.2 (on uniqueness of the Laplace transform). *If two continuous functions $f(t)$ and $\varphi(t)$ have the same transform $F(p)$, then they are identically equal.*

Theorem 28.3 (on linearity). *If $f(t)$ and $\varphi(t)$ are inverse transforms, then for any complex constants α and β*

$$\alpha f(t) + \beta \varphi(t) \risingdotseq \alpha F(p) + \beta \Phi(p). \tag{28.7}$$

This proposition follows from the linearity property of the integral that defines the transform

$$\int_0^{+\infty} (\alpha f(t) + \beta \varphi(t)) e^{-pt} dt = \alpha \int_0^{+\infty} f(t) e^{-pt} dt$$
$$+ \beta \int_0^{+\infty} \varphi(t) e^{-pt} dt = \alpha F(p) + \beta \Phi(p)$$

($\mathrm{Re}\, p > \max\{s_0, s_1\}$, where s_0, s_1 are infimums for the functions $f(t)$ and $\varphi(t)$, respectively).

From this property we obtain

$$\sin \omega t = \frac{e^{i\omega t} - e^{-i\omega t}}{2i} \doteqdot \frac{1}{2i}\left(\frac{1}{p - i\omega} - \frac{1}{p + i\omega}\right) = \frac{\omega}{p^2 + \omega^2},$$

i.e.,

$$\sin \omega t \doteqdot \frac{\omega}{p^2 + \omega^2}. \tag{28.8}$$

Likewise, we find that

$$\cos \omega t \doteqdot \frac{p}{p^2 + \omega^2}. \tag{28.9}$$

Further,

$$\sinh t = \frac{e^t - e^{-t}}{2} \doteqdot \frac{1}{2}\left(\frac{1}{p - 1} - \frac{1}{p + 1}\right) = \frac{1}{p^2 - 1}, \tag{28.10}$$

$$\cosh t \doteqdot \frac{p}{p^2 - 1}. \tag{28.11}$$

Theorem 28.4 (on similarity). *If $f(t)$ is an inverse transform and $F(p)$ is its Laplace transform, then for any constant $\alpha > 0$*

$$f(\alpha t) \doteqdot \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right). \tag{28.12}$$

◄ Setting $\alpha t = \tau$, we have

$$f(\alpha t) \doteqdot \int_0^{+\infty} f(\alpha t) e^{-pt} dt = \frac{1}{\alpha} \int_0^{+\infty} f(\tau) e^{-\frac{p}{\alpha}\tau} d\tau = \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right).$$

Using this theorem we deduce from (28.10) and (28.11)

$$\sinh \omega t \doteqdot \frac{\omega}{p^2 - \omega^2},$$

$$\cosh \omega t \doteqdot \frac{p}{p^2 - \omega^2}.$$

Theorem 28.5 (on differentiation of inverse transform). *Let $f(t)$ be an inverse transform of $F(p)$, and let $f'(t), f''(t), \ldots, f^{(n)}(t)$ be also inverse transforms, and $\bar{s} = \max\{s_0, s_1, \ldots, s_n\}$, where s_k is the infimum of $f^{(k)}(t)$ $(k = 0, 1, \ldots, n)$. Then*

$$f'(t) \doteqdot p\, F(p) - f(0) \tag{28.13}$$

and, in general,

$$f^{(n)}(t) \risingdotseq p^n F(p) - p^{n-1} f(0) - p^{n-2} f'(0) - \ldots - f^{(n-1)}(0). \tag{28.14}$$

Here $f^{(k)}(0)$ $(k = 0, 1, \ldots, (n-1))$ stands for the right limiting value $f^{(k)}(t)$:

$$f^{(k)}(0) = \lim_{t \to 0+0} f^{(k)}(t).$$

◄ Let $f(t) \risingdotseq F(p)$. We want to find the image of $f'(t)$. We have

$$f'(t) \risingdotseq \int_0^{+\infty} f'(t) e^{-pt} dt.$$

Integrating by parts gives

$$f'(t) \risingdotseq \int_0^{\infty} f'(t) e^{-pt} dt = (f(t) e^{-pt})\Big|_{t=0}^{t=+\infty} + p \int_0^{+\infty} f(t) e^{-pt} dt \tag{28.15}$$

The first term on the right vanishes at $t \to +\infty$, since for $\operatorname{Re} p = s > \bar{s}$ we have $|f(t)e^{-pt}| \leqslant Me^{-(s-\bar{s})t} \underset{t \to +\infty}{\longrightarrow} 0$; substituting $t = 0$ gives $-f(0)$.

The second term in (28.15) is equal to $p\,F(p)$. And so relation (28.15) becomes

$$f'(t) \risingdotseq p\,F(p) - f(0),$$

which proves formula (28.13). Specifically, if $f(0) = 0$, then $f'(t) \risingdotseq p\,F(p)$.

To find the Laplace transform of $f^{(n)}(t)$, we write

$$f^{(n)}(t) \risingdotseq \int_0^{+\infty} f^{(n)}(t) e^{-pt} dt,$$

whence, integrating by parts n times, we get

$$f^{(n)}(t) \risingdotseq p^n F(p) - p^{n-1} f(0) - p^{n+2} f'(0) - \ldots - f^{(n-1)}(0). ►$$

Example. Using the theorem on differentiability of the inverse transform, find the Laplace transform of the function $f(t) = \sin^2 t$.

◄ Let $f(t) \risingdotseq F(p)$. Then

$$f'(t) \risingdotseq p\,F(p) - f(0).$$

But $f(0) = 0$, and $f'(t) = 2 \sin t \cos t = \sin 2t \risingdotseq 2/(p^2 + 4)$. Therefore, $2/(p^2 + 4) = p\,F(p)$, hence $F(p) = 2/p(p^2 + 4) \fallingdotseq \sin^2 t$. ►

Theorem 28.5 establishes the remarkable property of the integral Laplace transform: it, like the Fourier transform, converts differentiation into the algebraic operation of multiplication by p.

Inclusion formula. If $f(t)$ and $f'(t)$ are inverse transforms, then

$$\lim_{\mathrm{Re}\, p \to +\infty} p\, F(p) = f(0). \tag{28.16}$$

Really, $f'(t) \risingdotseq p\, F(p) - f(0)$. By Corollary of Theorem 28.1, any transform tends to zero as $\mathrm{Re}\, p = s \to +\infty$. Therefore $\lim\limits_{\mathrm{Re}\, p \to +\infty} [p\, F(p) - f(0)] = 0$, hence we obtain (28.16).

Theorem 28.6 (on transform differentiation). *Differentiation of a transform reduces to multiplication of the original function by $-t$. In general,*

$$F^{(n)}(p) \fallingdotseq (-1)^n t^n f(t). \tag{28.17}$$

◄ Since $F(p)$ in the half-plane $\mathrm{Re}\, p = s > s_0$ is an analytic function, it can be differentiated with respect to p to yield

$$F'(p) = -\int_0^{+\infty} t f(t) e^{-pt} dt, \quad F''(p) = \int_0^{\infty} t^2 f(t) e^{-pt} dt, \ \ldots,$$

$$F^{(n)}(p) = \int_0^{+\infty} (-1)^n t^n f(t) e^{-pt} dt.$$

This exactly means that $F^{(n)}(p) \fallingdotseq (-1)^n t^n f(t)$. ►

Example. Using Theorem 28.6 find the transform of the function $\varphi(t) = t^n$.

◄ It is well known that $1 \risingdotseq 1/p$. Here $f(t) = 1$, $F(p) = 1/p$. Hence $(1/p)' \fallingdotseq (-t)\cdot 1$ or $1/p^2 \fallingdotseq t$. Again using the theorem we find that $(1/p^2)' \fallingdotseq (-t)\cdot t$ or $(1 \times 2)/p^3 \fallingdotseq t^2$. In general $t^n \risingdotseq n!/p^{n+1}$.

Theorem 28.7 (on integration of inverse transform). *Integration of an inverse transform reduces to division of the transform by p: if $f(t) \risingdotseq F(p)$, then*

$$\int_0^t f(t)\, dt \risingdotseq \frac{F(p)}{p}. \tag{28.18}$$

◄ We put

$$\varphi(t) = \int_0^t f(t)\, dt. \tag{28.19}$$

It can be readily verified that if $f(t)$ is an inverse transform, $\varphi(t)$ is also an inverse transform and $\varphi(0) = 0$. Let $\varphi(t) \risingdotseq \Phi(p)$. By (28.19)

$$f(t) = \varphi'(t) \risingdotseq p\, \Phi(p) - \varphi(0) = p\, \Phi(p),$$

so that $f(t) \risingdotseq p\, \Phi(p)$. On the other hand, $f(t) \risingdotseq F(p)$, hence $F(p) = p\, \Phi(p)$, i.e., $\Phi(p) = F(p)/p$. This is equivalent to the relation (28.18) we seek to prove.

Example. Find the transform of the function

$$\varphi(t) = \int_0^t \cos t\, dt.$$

◄ Here $f(t) = \cos t$, so that $F(p) = p/(p^2 + 1)$. Therefore,

$$\int_0^t \cos t\, dt \risingdotseq \frac{1}{p^2 + 1}. \text{ ►}$$

Theorem 28.8 (on integration of transform). *If the integral* $\int_p^\infty F(p)dp$ *converges, then it is the transform of the function* $f(t)/t$, *i.e.*,

$$\frac{f(t)}{t} \risingdotseq \int_0^\infty F(p)dp. \tag{28.20}$$

◄ Indeed,

$$\int_p^\infty F(p)dp = \int_p^\infty \left\{ \int_0^{+\infty} f(t)e^{-pt}dt \right\} dp.$$

Assuming that the integration path (p, ∞) lies in the half-plane $\operatorname{Re} p \geqslant a > s_0$, we can change the order of integration for $t > 0$

$$\int_p^\infty F(p)dp = \int_p^\infty \left\{ \int_0^{+\infty} f(t)e^{-pt}dt \right\} dp$$

$$= \int_0^{+\infty} f(t) \left\{ \int_p^\infty e^{-pt}dp \right\} dt = \int_0^{+\infty} \frac{f(t)}{t} e^{-pt}dt.$$

This means that $\int_p^\infty F(p)dp$ is the Laplace transform of $f(t)/t$. ►

Example. Find the transform for the function $\dfrac{\sin t}{t}$.

◄ It is well known that $\sin t \risingdotseq 1/(p^2 + 1)$. Therefore,

$$\frac{\sin t}{t} \risingdotseq \int_0^\infty \frac{dp}{p^2 + 1} = \tan^{-1} p\Big|_p^\infty = \frac{\pi}{2} - \tan^{-1} p = \cot^{-1} p. \text{ ►}$$

Theorem 28.9 (on delay). *If* $f(t) \risingdotseq F(p)$, *then for any positive* τ *(delay)*

$$f(t - \tau) \risingdotseq e^{-p\tau}F(p).$$

◄ Since $f(t-\tau) \equiv 0$ for $t < \tau$ (Fig. 28.4), we have

$$f(t-\tau) \doteqdot \int_0^{+\infty} f(t-\tau)e^{-pt}dt = \int_\tau^{+\infty} f(t-\tau)e^{-pt}dt. \tag{28.21}$$

Putting $\xi = t - \tau$, we find $dt = d\xi$. At $t = \tau$ we obtain $\xi = 0$; when $t = +\infty$ we will have $\xi = +\infty$.

Therefore, relation (28.21) becomes

$$f(t-\tau) \doteqdot \int_0^{+\infty} f(\xi)e^{-p(\xi+\tau)}d\xi = e^{-p\tau}\int_0^{+\infty} f(\xi)e^{-p\xi}d\xi = e^{-p\tau}F(p),$$

i.e., we indeed have $f(t-\tau) \doteqdot e^{-p\tau}F(p)$. ►

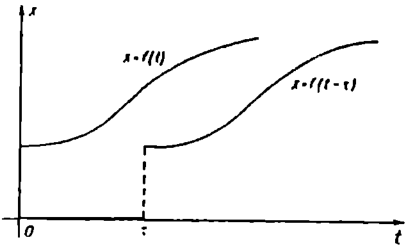


Fig. 28.4

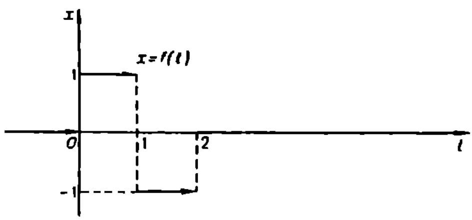


Fig. 28.5

Example. Find the transform for the function $f(t)$ presented graphically in Fig. 28.5.

◄ We write $f(t)$ in the form

$$f(t) = \eta(t) - 2\eta(t-1) + \eta(t-2).$$

This expression can be derived as follows. Consider a function $f_1(t) = \eta(t)$ for $t \geqslant 0$ (Fig. 28.6*a*). We subtract from it the function $f_2(t) = 2\eta(t-1) = \begin{cases} 0 & \text{for } t < 1 \\ 2 & \text{for } t \geqslant 1 \end{cases}$. The difference $f_1(t) - f_2(t)$ will be equal to unity for $t \in [0, 1)$ and to -1 for $t \geqslant 1$ (Fig. 28.6*b*). To the difference we add the function $f_3(t) = \eta(t-2) = \begin{cases} 0 & \text{for } t < 2 \\ 1 & \text{for } t \geqslant 2 \end{cases}$. As a result, we obtain $f(t)$ (Fig. 28.6*c*) so that

$$f(t) = \eta(t) - 2\eta(t-1) + \eta(t-2).$$

Using the delay theorem we obtain from this

$$F(p) = \frac{1}{p} - \frac{2}{p}e^{-p} + \frac{1}{p}e^{-2p}. \;►$$

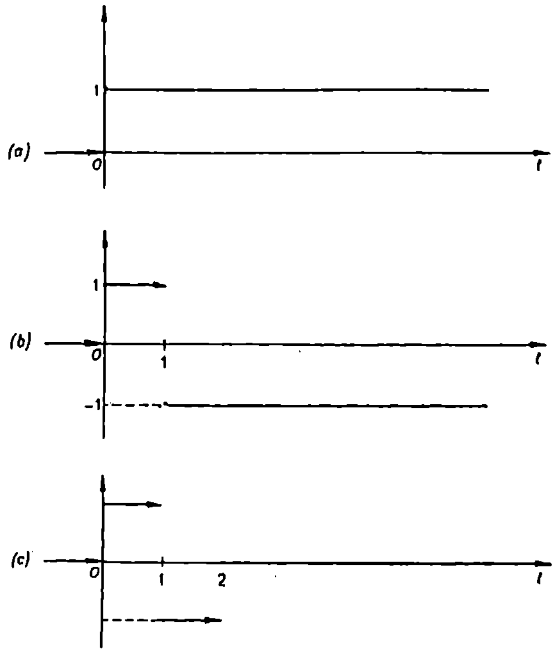


Fig. 28.6

Theorem 28.10 (on displacement). *If* $f(t) \risingdotseq F(p)$, *then for any complex number* p_0

$$e^{p_0 t} f(t) \risingdotseq F(p - p_0).$$

◄ Indeed,

$$e^{p_0 t} f(t) \risingdotseq \int_0^{+\infty} e^{p_0 t} f(t) e^{-pt}\, dt = \int_0^{+\infty} f(t) e^{-(p-p_0)t}\, dt = F(p - p_0). \text{ ►}$$

The theorem enables us from known transforms of functions to determine the transforms of the functions multiplied by $e^{p_0 t}$, e.g.,

$$\sin \omega t \risingdotseq \frac{\omega}{p^2 + \omega^2}, \quad \text{hence} \quad e^{-\lambda t} \sin \omega t \risingdotseq \frac{\omega}{(p + \lambda)^2 + \omega^2}$$

$$\cos \omega t \risingdotseq \frac{p}{p^2 + \omega^2}, \quad \text{hence} \quad e^{\lambda t} \cos \omega t \risingdotseq \frac{p - \lambda}{(p - \lambda)^2 + \omega^2}.$$

Convolution of functions. Multiplication theorem. It will be recalled that if functions $f(t)$ and $\varphi(t)$ are defined and continuous for all t, then we define the *convolution* $(f * \varphi)(t)$ of the functions as a new function of

t given by

$$(f * \varphi)(t) = \int_{-\infty}^{+\infty} f(\tau)\varphi(t - \tau)d\tau$$

(if the integral exists).

For inverse transforms $f(t)$ and $\varphi(t)$ the convolution operation is always possible, and

$$(f * \varphi)(t) = \int_0^t f(\tau)\varphi(t - \tau)d\tau. \tag{28.22}$$

◄ The product of the inverse transforms $f(\tau)\varphi(t - \tau)$ as a function of τ is a finite function, i.e., it vanishes outside of some finite interval (here outside of the interval $0 \leqslant \tau \leqslant t$). For finite continuous functions convolution operation is possible, and so we will arrive at (28.22). ►

It can readily be verified that

$$\int_0^t f(\tau)\varphi(t - \tau)d\tau = \int_0^t \varphi(\tau)f(t - \tau)d\tau,$$

i.e., that the convolution operation is *commutative*.

Theorem 28.11 (on multiplication). *If* $f(t) \risingdotseq F(p)$, $\varphi(t) \risingdotseq \Phi(p)$, *then the convolution* $(f * \varphi)(t)$ *has the transform* $F(p)\Phi(p)$:

$$\int_0^t f(\tau)\varphi(t - \tau)d\tau \risingdotseq F(p)\Phi(p),$$

or

$$(f * \varphi)(t) \risingdotseq F(p)\Phi(p).$$

◄ We can show that the convolution $(f * \varphi)(t)$ of inverse transforms is an inverse transform with the infimum $s^* = \max\{s_1, s_2\}$ where s_1, s_2 are the infimums of $f(t)$ and $\varphi(t)$, respectively. We now find the transform of the convolution

$$\int_0^t f(\tau)\varphi(t - \tau)d\tau \risingdotseq \int_0^{+\infty} e^{-pt}\left\{\int_0^t f(\tau)\varphi(t - \tau)d\tau\right\}dt. \tag{28.23}$$

Making use of the fact that

$$\int_0^t f(\tau)\varphi(t - \tau)d\tau = \int_0^{+\infty} f(\tau)\varphi(t - \tau)d\tau,$$

$$\varphi(t - \tau) \equiv 0 \quad \text{for} \quad \tau > t,$$

we will obtain

$$\int_0^{+\infty} e^{-pt}\left\{\int_0^t f(\tau)\varphi(t-\tau)d\tau\right\}dt$$

$$= \int_0^{+\infty} e^{-pt}\left\{\int_0^{+\infty} f(\tau)\varphi(t-\tau)d\tau\right\}dt.$$

Changing the order of integration in the integral on the right-hand side (for Re $p = s > s^*$ the operation is legitimate) and using the delay theorem, we get

$$\int_0^{+\infty} e^{-pt}\left\{\int_0^t f(\tau)\varphi(t-\tau)d\tau\right\}dt$$

$$= \int_0^{+\infty} f(\tau)\left\{\int_0^{+\infty} e^{-pt}\varphi(t-\tau)dt\right\}d\tau$$

$$= \int_0^{\infty} f(\tau)e^{-p\tau}\Phi(p)d\tau = \Phi(p)F(p). \tag{28.24}$$

From (28.23) and (28.24) we thus find

$$\int_0^t f(\tau)\varphi(t-\tau)d\tau \risingdotseq F(p)\Phi(p), \tag{28.25}$$

i.e., *multiplication of transforms corresponds to convolution of the inverse transforms*

$$F(p)\Phi(p) \fallingdotseq (f*\varphi)(t).$$

Example. Find the transform for the function

$$\psi(t) = \int_0^t \tau \sin(t-\tau)d\tau.$$

◄ The function $\psi(t)$ is the convolution of the functions $f(t) = t$ and $\varphi(t) = \sin t$. By the multiplication theorem,

$$\psi(t) \risingdotseq F(p)\Phi(p) = \frac{1}{p^2(p^2+1)}. \text{ ►}$$

Problem. Let a function $f(t)$ with period T be an inverse transform. Show that its Laplace transform $F(p)$ will be

$$F(p) = \frac{1}{1-e^{-pT}}\int_0^T f(t)e^{-pt}dt \quad (\text{Re } p = s > 0).$$

28.3 Inverse Transform

Let us formulate the problem: given a function $F(p)$, find the function $f(t)$ whose transform is $F(p)$.

We will now formulate conditions that are sufficient for a function $F(p)$ of a complex variable p to be a transform.

Theorem 28.12. *If a function $F(p)$ that is analytic in the half-plane* $\operatorname{Re} p = s > s_0$

(1) *tends to zero as $|p| \to +\infty$ in any half-plane* $\operatorname{Re} p \geqslant a > s_0$ *uniformly in* $\arg p$,

(2) *the integral* $\int\limits_{a-i\infty}^{a+i\infty} F(p)\,dp$ $(a > s_0)$ *converges absolutely,*

then $F(p)$ is the Laplace transform of some inverse transform $f(t)$.

Problem. Can the function $F(p) = (p + 1)/p$ be the transform of some function?

We will describe some of the methods of recovering the original function (inverse transform) from the Laplace transform.

(1) Recovering the inverse transform using tables of transforms. There can be found in the literature fairly extensive tables of Laplace transforms of various functions.

To begin with, one attempts to reduce a function $F(p)$ to a simpler, "table" form, e.g., when $F(p)$ is a fractional rational function of the argument p, one expands it into elementary fractions and uses suitable properties of the Laplace transform.

Examples. (1) Find the inverse transform for $F(p) = p/(p^2 + 2p + 5)$.

◄ We write $F(p)$ in the form

$$F(p) = \frac{p}{p^2 + 2p + 5} = \frac{p}{(p+1)^2 + 2^2}$$
$$= \frac{p+1}{(p+1)^2 + 2^2} - \frac{1}{2}\,\frac{2}{(p+1)^2 + 2^2}\,.$$

Using the displacement theorem and the linearity property for the Laplace transforms, we obtain

$$f(t) = e^{-t}\cos 2t - \frac{1}{2}\,e^{-t}\sin 2t. \ ►$$

(2) Find the inverse transform for $F(p) = \dfrac{1}{p^2(p^2+1)}$.

◄ We write $F(p)$ in the form

$$F(p) = \frac{1}{p^2(p^2+1)} = \frac{1}{p^2} - \frac{1}{p^2+1}\,.$$

Hence $f(t) = t - \sin t$. ►

Use of inversion theorem and its corollaries. Theorem 28.13 (inversion theorem). *If a function $f(t)$ is an inverse transform with the infimum s_0 and $F(p)$ is its transform, then at any point where $f(t)$ is continuous, it will be*

$$f(t) = \frac{1}{2\pi i} \int_{s-i\infty}^{s+i\infty} F(p)e^{pt}\,dp, \tag{28.26}$$

where the integration is along any straight line Re $p = s =$ const $> s_0$ *and the integral is understood to be the principal value*, i.e.,

$$\lim_{N\to+\infty} \int_{s-iN}^{s+iN} F(p)e^{pt}\,dp.$$

Formula (28.26) is known as the *inversion formula* for the Laplace transform or as the *Mellin formula.*

◄ Suppose that $f(t)$ is a piecewise smooth function on each finite interval $[0, a]$ with the infimum s_0. Consider the function $\varphi(t) = f(t)e^{-st}$, where s is any real number such that $s > s_0$.

The function $\varphi(t)$ meets the conditions for the Fourier formula to be applicable, and so holds the inversion formula for the Fourier transform

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Phi(\xi)e^{i\xi t}\,d\xi, \tag{28.27}$$

where

$$\Phi(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(t)e^{-i\xi t}\,dt = \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} \varphi(t)e^{-i\xi t}\,dt. \tag{28.28}$$

Here $\varphi(t) \equiv 0$ for $t < 0$.

Substituting the expression $\varphi(t) = f(t)e^{-st}$ into (28.28) gives

$$\Phi(\xi) = \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-i\xi t}e^{-st}f(t)\,dt$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-(s+i\xi)t}f(t)\,dt = \frac{1}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-pt}f(t)\,dt = \frac{1}{\sqrt{2\pi}} F(p),$$

where $F(p)$ is the Laplace transform for $f(t)$ when $p = s + i\xi$.

Formula (28.27) can be rewritten as

$$\varphi(t) = f(t)e^{-st} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\xi t}F(s + i\xi)\, d\xi.$$

From this we derive the inversion formula for the Laplace transform

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{(s+i\xi)t}F(s + i\xi)d\xi = \frac{1}{2\pi i}\int_{s-i\infty}^{s+i\infty} e^{pt}F(p)dp, \quad s > s_0. \blacktriangleright$$

As a corollary of the inversion theorem we deduce the uniqueness theorem.

Theorem 28.14. *Two continuous functions $f(t)$ and $\varphi(t)$ with the same transform $F(p)$ are identical.*

Direct calculation of the inversion integral (28.26) is usually a hard exercise. But with certain additional constraints on $F(p)$ recovering the inverse transform from its transform is much simpler.

Theorem 28.15. *Let a transform $F(p)$ be a fractional rational function with poles $p_1, p_2, \ldots, p_n$. The inverse transform for $F(p)$ will then be the function $f(t)\eta(t)$, where*

$$f(t) = \sum_{k=1}^{n} \operatorname*{res}_{p=p_k} (F(p)e^{pt}). \tag{28.29}$$

$\blacktriangleleft$ Let the transform $F(p)$ be a fractional rational function, i.e., $F(p) = A(p)/B(p)$, where $A(p)$, $B(p)$ are polynomials in p (mutually simple). The degree of the numerator must be smaller than that of the denominator, since for any transform we have

$$\lim_{\operatorname{Re} p \to +\infty} F(p) = 0,$$

whereas for rational transforms $\lim\limits_{p\to\infty} F(p) = 0$, whatever the manner in which the point p tends to infinity.

Let the roots of $B(p)$, which are poles of $F(p)$, be $p_1, p_2, \ldots, p_n$, and their multiplicities be $r_1, r_2, \ldots, r_n$, respectively.

If the number s in the inversion formula (28.26) is larger than all Re p_k ($k = 1, 2, \ldots, n$), then by the inversion formula, which is applicable under these conditions, we will get

$$f(t) = \frac{1}{2\pi i}\int_{s-i\infty}^{s+i\infty} F(p)e^{pt}dp.$$

Consider the closed contour Γ_R (Fig. 28.7), which is traced in the positive direction and which consists of the arc C_R of a circle of radius R with centre at the origin, and the chord AB that subtends it, which is a segment of the straight line $\operatorname{Re} p = s = \text{const}$, the radius R being so large that all the poles of $F(p)$ lie inside Γ_R.

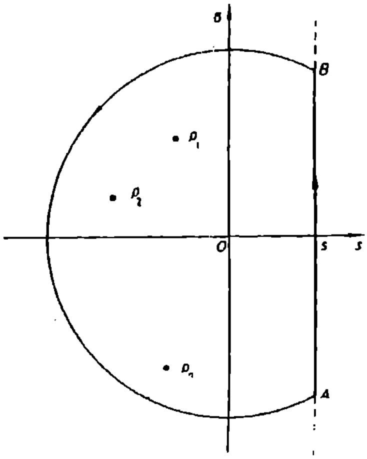


Fig. 28.7

By the Cauchy theorem on residues, for any R satisfying the above condition we will have

$$\frac{1}{2\pi i}\int_{\Gamma_R} F(p)e^{pt}dp = \sum_{k=1}^{n} \operatorname*{res}_{p=p_k} (F(p)e^{pt}),$$

or

$$\frac{1}{2\pi i}\int_{AB} F(p)e^{pt}dp + \frac{1}{2\pi i}\int_{C_R} F(p)e^{pt}dp = \sum_{k=1}^{n} \operatorname*{res}_{p=p_k} (F(p)e^{pt}). \quad (28.30)$$

The second term on the left of (28.30) tends to zero as $R \to \infty$. This follows from the Jordan lemma, if we replace iz by p and take into account that $F(p) \to 0$ as $p \to \infty$.

If in (28.30) we pass to the limit as $R \to \infty$, we will obtain on the left

$$f(t) = \frac{1}{2\pi i}\int_{s-i\infty}^{s+i\infty} F(p)e^{pt}dp$$

and on the right the sum of residues over all the poles of $F(p)$, i.e.,

$$f(t) = \sum_{k=1}^{n} \operatorname*{res}_{p=p_k} (F(p)e^{pt}). \blacktriangleright$$

We will make use of the formula to compute residues (see Chap. 26) to find that

$$f(t) = \sum_{k=1}^{n} \frac{1}{(r_k - 1)!} \lim_{p \to p_k} \frac{d^{r_k - 1}}{dp^{r_k - 1}} \left\{ \frac{A(p)}{B(p)} e^{pt} (p - p_k)^{r_k} \right\}. \quad (28.31)$$

If all the poles $p_1, p_2, \ldots, p_n$ are simple, then $\operatorname*{res}_{p=p_k} (F(p)e^{pt}) = A(p_k)/B'(p_k)e^{p_k t}$ $(k = 1, 2, \ldots, n)$, and (28.31) becomes

$$f(t) = \sum_{k=1}^{n} \frac{A(p_k)}{B'(p_k)} e^{p_k t}. \quad (28.32)$$

Example. Find the inverse transform for the function $F(p) = 1/(p^2 + 1)$.

◄ Function $F(p)$ has simple poles $p_1 = i$, $p_2 = -i$. Using formula (28.32), we find

$$f(t) = \operatorname*{res}_{p=i} \left\{ \frac{e^{pt}}{p^2 + 1} \right\} + \operatorname*{res}_{p=-i} \left\{ \frac{e^{pt}}{p^2 + 1} \right\} = \frac{e^{it}}{2i} + \frac{e^{-it}}{-2i}$$

$$= \frac{e^{it} - e^{-it}}{2i} = \sin t. \blacktriangleright$$

Theorem 28.16. *Let a transform $F(p)$ be an analytic function at the point $p = \infty$. Suppose that its expansion in the neighbourhood $|p| > R$ of the infinite point has the form*

$$F(p) = \frac{c_1}{p} + \frac{c_2}{p^2} + \ldots + \frac{c_n}{p^n} + \ldots = \sum_{k=1}^{\infty} \frac{c_k}{p^k}. \quad (28.33)$$

The inverse transform for $F(p)$ will then be the function $f(t)\eta(t)$, where

$$f(t) = \sum_{k=1}^{\infty} \frac{c_k}{(k-1)!} t^{k-1}. \quad (28.34)$$

By way of example, we again take $F(p) = 1/(p^2 + 1)$. We have

$$\frac{1}{p^2 + 1} = \frac{1}{p^2 \left(1 + \frac{1}{p^2}\right)} = \sum_{n=0}^{\infty} \frac{(-1)^n}{p^{2n+2}}$$

so that

$$f(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1} = \sin t.$$

28.4 Applications of Laplace Transform (Operational Calculus)

Solution of linear differential equations with constant coefficients. Suppose that we are given a linear differential equation with constant coefficients

$$a_0 x''(t) + a_1 x'(t) + a_2 x(t) = f(t), \tag{28.35}$$

where a_0, a_1, a_2 are real numbers. We want to find the solution $x(t)$ of equation (28.35) for $t > 0$, such that it obeys the initial conditions

$$x(0) = x_0, \quad x'(0) = x_1. \tag{28.36}$$

We will take $f(t)$ to be the inverse transform, such that

$$f(t) \risingdotseq F(p), \quad x(t) \risingdotseq X(p).$$

By the theorem on differentiation of the inverse transform,

$$x'(t) \risingdotseq pX(p) - x_0,$$
$$x''(t) \risingdotseq p^2 X(p) - px_0 - x_1.$$

Using the property of linearity of the Laplace transform and Theorem 28.5 we pass in (28.35) from inverse transforms to transforms

$$a_0 p^2 X(p) - a_0 p x_0 - a_0 x_1 + a_1 p X(p) - a_1 x_0 + a_2 X(p) = F(p)$$

or

$$(a_0 p^2 + a_1 p + a_2) X(p) = F(p) + a_0 x_0 p + a_0 x_1 + a_1 x_0. \tag{28.37}$$

But this is no longer a differential equation; instead, this is an algebraic equation in the transform $X(p)$ of the original function. It is called an *operator equation.* Solving this gives

$$X(p) = \frac{F(p) + a_0 x_0 p + a_0 x_1 + a_1 x_0}{a_0 p^2 + a_1 p + a_2},$$

which is the operator solution of the problem (28.35-36).

The inverse transform for $X(p)$ will be the required solution $x(t)$ of the problem.

The general case of the linear differential equation of order n ($n \geqslant 1$) with constant coefficients is essentially the same as for the case $n = 2$.

We will now provide the general scheme of solution of the Cauchy problem

$$
\begin{tikzcd}
\text{I} & \boxed{\text{Cauchy problem in the space of inverse transforms}} \arrow[d, "\mathscr{L}"] & \boxed{\text{Solution of the Cauchy problem}} & \text{IV} \\
\text{II} & \boxed{\text{Operator equation in the space transforms}} \arrow[r] & \boxed{\text{Solution of the operator equation}} \arrow[u, "\mathscr{L}^{-1}"'] & \text{III}
\end{tikzcd}
$$

Here $\mathscr{L}$ is application to I of the Laplace transform, $\mathscr{L}^{-1}$ is application to III of the inverse Laplace transform.

Example. Solve the Cauchy problem

$$x''(t) + x(t) = 2\cos t,$$
$$x(0) = 0, \quad x'(0) = 1.$$

◄ Here $f(t) = 2\cos t \doteqdot 2p/(p^2 + 1)$. Let $x(t) \doteqdot X(p)$. Then $x''(t) \doteqdot p^2X(p) - 1$.

The operator equation is

$$p^2X(p) - 1 + X(p) = \frac{2p}{p^2 + 1},$$

hence

$$X(p) = \frac{2p}{(p^2 + 1)^2} + \frac{1}{p^2 + 1}.$$

By the theorem on differentiation of transforms

$$\frac{2p}{(p^2 + 1)^2} = -\left(\frac{1}{p^2 + 1}\right)_p' \doteqdot t\sin t.$$

Therefore,

$$x(t) = t\sin t + \sin t. \ ►$$

Duhamel formula. In applications of operational calculus to solutions of differential equations one often uses the consequence of the multiplication theorem, known as the *Duhamel formula*.

Let $f(t)$ and $\varphi(t)$ be inverse transforms, and $f(t)$ be continuous on the interval $[0, +\infty)$, and $\varphi(t)$ be continuously differentiable on $[0, +\infty)$. If

$f(t) \risingdotseq F(p)$ and $\varphi(t) \risingdotseq \Phi(p)$, then by the multiplication theorem

$$\psi(t) = \int_0^t f(\tau)\varphi(t - \tau)\,d\tau \risingdotseq F(p)\Phi(p).$$

It can easily be verified that the function $\psi(t)$ is continuously differentiable on $[0, +\infty)$, and

$$\psi'(t) = \frac{d}{dt}\left(\int_0^t f(\tau)\varphi(t - \tau)\,d\tau\right) = f(t)\varphi(0) + \int_0^t f(\tau)\varphi'(t - \tau)\,d\tau.$$

By virtue of the differentiation rule for inverse transforms and considering that $\psi(0) = 0$, we will obtain the Duhamel formula

$$f(t)\varphi(0) + \int_0^t f(\tau)\varphi'(t - \tau)\,d\tau \risingdotseq p\,F(p)\Phi(p). \tag{28.38}$$

We will now illustrate the applications of the formula. Suppose that we wish to solve a linear differential equation of order n ($n \geqslant 1$) with constant coefficients

$$L[x(t)] = f(t) \tag{28.39}$$

with zero initial conditions

$$x(0) = x'(0) = \ldots = x^{(n-1)}(0) = 0 \tag{28.40}$$

(the last limitation is of no significance: a problem with nonzero initial conditions can be reduced to a problem with zero conditions by a change of the desired function).

If we know the solution $x_1(t)$ of the differential equation

$$L[x(t)] = 1 \tag{28.41}$$

with the same left-hand side and the right-hand side equal to unity, also with zero initial conditions

$$x_1(0) = x_1'(0) = \ldots = x_1^{(n-1)}(0) = 0, \tag{28.42}$$

then the Duhamel formula (28.38) allows us directly to obtain the solution of the original problem (28.39-40).

The operator equations for the problems (28.39-40) and (28.41-42) will be, respectively,

$$A(p)X(p) \risingdotseq F(p), \tag{28.43}$$

and

$$A(p)X_1(p) = \frac{1}{p}, \tag{28.44}$$

where $F(p)$ is the transform of $f(t)$. From (28.43) and (28.44) we readily find

$$X(p) = p\,X_1(p)F(p).$$

From this, by the Duhamel formula, we have

$$x(t) = f(t)x_1(0) + \int_0^t f(\tau)x_1'(t-\tau)\,d\tau$$

or, since $x_1(0) = 0$,

$$x(t) = \int_0^t f(\tau)x_1'(t-\tau)\,d\tau. \tag{28.45}$$

Example. Solve the Cauchy problem

$$x''(t) - x(t) = \frac{1}{1+e^t},$$

$$x(0) = x'(0) = 0.$$

◄ Consider the auxiliary problem

$$x_1''(t) - x_1(t) = 1,$$
$$x_1(0) = x_1'(0) = 0.$$

Using the operator technique, we find that

$$X_1(p) = \frac{1}{p(p^2-1)} \doteqdot \int_0^t \sinh t\,dt = x_1(t).$$

From (28.45), we get the solution $x(t)$ of the original problem

$$x(t) = \int_0^t \frac{1}{1+e^\tau}\sinh(t-\tau)\,d\tau = \frac{1}{2}(e^t - te^t - 1)$$

$$+ \sinh t \ln\frac{1+e^t}{2}. \text{ ►}$$

Integration of systems of linear differential equations with constant coefficients. The integration occurs as in solving one linear differential equation, namely, by passing from a system of differential equation to a system of operator equations. Solving the latter as a system of linear algebraic equations for the transforms of the desired functions, we will get the operator solution of the system. Its inverse transform will be the solution of the original system of differential equations.

Example. Find the solution of the linear system

$$\begin{cases} x'(t) = -y(t), \\ y'(t) = 2x(t) + 2y(t), \end{cases}$$

the initial conditions being $x(0) = y(0) = 1$.

Let $x(t) \risingdotseq X(p)$, $y(t) \risingdotseq Y(p)$. Using the linearity property of the Laplace transform and the theorem on differentiation of inverse transforms, we reduce the original Cauchy problem to the operator system

$$\begin{cases} pX(p) - 1 = -Y(p), \\ pY(p) - 1 = 2X(p) + 2Y(p). \end{cases}$$

Solving the latter for $X(p)$ and $Y(p)$ gives

$$X(p) = \frac{p-3}{p^2 - 2p + 2} = \frac{p-1}{(p-1)^2 + 1} - 2\frac{1}{(p-1)^2 + 1}$$

$$\risingdotseq e^t \cos t - 2e^t \sin t,$$

$$Y(p) = \frac{p+2}{p^2 - 2p + 2} = \frac{p-1}{(p-1)^2 + 1} + \frac{3}{(p-1)^2 + 1}$$

$$\risingdotseq e^t \cos t + 3e^t \sin t.$$

And so the original Cauchy problem has the solution

$$x(t) = e^t \cos t - 2\, e^t \sin t,$$

$$y(t) = e^t \cos t + 3\, e^t \sin t.$$

Solution of integral equations. It will be recalled that we define an integral equation as an equation in which the unknown function appears under the integration sign.

We will only consider an equation of the type

$$\varphi(t) = f(t) + \int_0^t K(t-\tau)\varphi(\tau)d\tau \tag{28.46}$$

which is called the linear *Volterra integral equation* of the second kind with kernel $K(t-\tau)$ that depends on the difference of arguments (convolution-type equation).

Here $\varphi(t)$ is the required function, $f(t)$ and $K(t)$ are specified functions.

Let $f(t)$ and $K(t)$ be inverse transforms, $f(t) \risingdotseq F(p)$, $K(t) \risingdotseq \mathscr{K}(p)$. Applying to both sides of (28.46) the Laplace transform and using the multiplication theorem, we will get

$$\Phi(p) = F(p) + \mathscr{K}(p)\Phi(p), \tag{28.47}$$

where $\Phi(p) \doteqdot \varphi(t)$. From (28.47),

$$\Phi(p) = \frac{F(p)}{1 - \mathscr{K}(p)} \quad (\mathscr{K}(p) \neq 1).$$

The inverse transform for $\Phi(p)$ will be the solution of the integral equation (28.46).

Example. Solve the integral equation

$$\varphi(t) = t + \int_0^t \sin(t - \tau)\varphi(\tau)d\tau. \tag{28.48}$$

◀ Applying the Laplace transform to both sides of (28.48), we will obtain

$$\Phi(p) = \frac{1}{p^2} + \frac{1}{p^2 + 1}\Phi(p),$$

hence $\Phi(p) = 1/p^2 + 1/p^4 \doteqdot t + t^3/6$.

The function $\varphi(t) = t + t^3/6$ is the solution of the equation (28.48) (substitution $\varphi(t) = t + t^3/6$ into (28.48) turns the latter into an identity in t). ▶

Remark. The Laplace transform can also be used in solutions of some problems of mathematical physics (see Chap. 30).

Table 28.1 Some Original Functions and their Laplace Transforms

Original function	Laplace transform
1	$\frac{1}{p}$
$t^n \quad (n = 1, 2, \ldots)$	$\frac{n!}{p^{n+1}}$
$t^\alpha \quad (\alpha > -1)$	$\frac{\Gamma(\alpha + 1)}{p^{n+1}}$
$e^\lambda \quad (\lambda = \alpha + i\beta)$	$\frac{1}{p - \lambda}$

Table 28.1 (concluded)

Original function	Laplace transform
$\sin \omega t$	$\dfrac{\omega}{p^2 + \omega^2}$
$\cos \omega t$	$\dfrac{p}{p^2 + \omega^2}$
$\sinh \omega t$	$\dfrac{\omega}{p^2 - \omega^2}$
$\cosh \omega t$	$\dfrac{p}{p^2 - \omega^2}$
$e^{\lambda t} \sin \omega t$	$\dfrac{\omega}{(p - \lambda)^2 + \omega^2}$
$e^{\lambda t} \cos \omega t$	$\dfrac{p - \lambda}{(p - \lambda)^2 + \omega^2}$
$\sin (t - \tau), \quad \tau > 0$	$\dfrac{e^{-p\tau}}{p^2 + 1}$
$\cos (t - \tau), \quad \tau > 0$	$\dfrac{e^{-p\tau}}{p^2 + 1}$
$J_n(t) \quad (n = 0, 1, \ldots)$	$\dfrac{(\sqrt{p^2 + 1} - p)^n}{\sqrt{p^2 + 1}}$
$\mathrm{Erf}\left(\dfrac{\alpha}{2\sqrt{t}}\right) = \dfrac{2}{\sqrt{\pi}} \displaystyle\int_{\alpha/2\sqrt{t}}^{+\infty} e^{-x^2} dx$	$\dfrac{1}{p} e^{-\alpha\sqrt{p}}$

Exercises

Which of the following functions are inverse transforms?

1. $f(t) = e^t \cos t$. 2. $f(t) = \dfrac{1}{t^2 + 1}$. 3. $f(t) = \dfrac{1}{(t-1)^2}$.

Using the properties of the Laplace transform find the transforms of the following functions:

4. $f(t) = \sin^2 t$. 5. $f(t) = \cos mt \cos nt$. 6. $f(t) = t \cos t$. 7. $f(t) = (t + 1) \sin 2t$. 8. $f(t) = e^{2t} \sin t$. 9. $f(t) = e^{-t} \cos 2t$.

Find the transforms of the functions specified graphically:

10. In Fig. 28.8*a*. 11. In Fig. 28.8*b*. 12. In Fig. 28.8*c*.

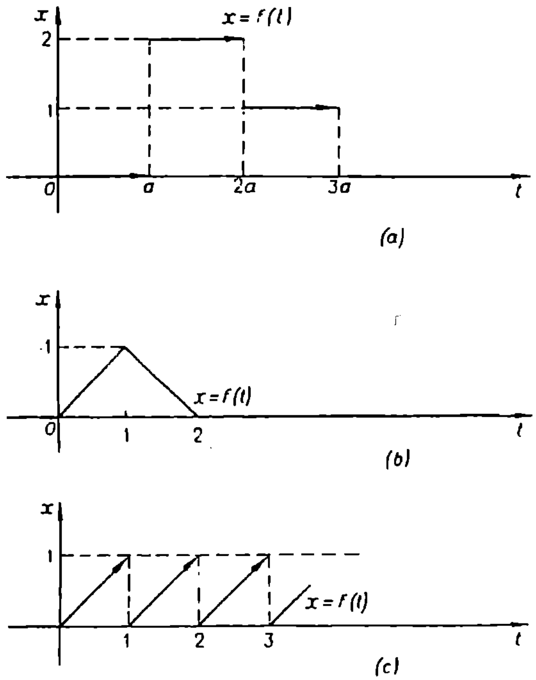


Fig. 28.8

Find the inverse transforms from the given transforms:

13. $F(p) = \dfrac{1}{p^2 + 4p + 5}$. 14. $F(p) = \dfrac{p}{(p+1)^2}$. 15. $F(p) = \dfrac{p}{(p^2+1)^2}$.

16. $F(p) = \dfrac{p}{p^2 - 2p + 3}$. 17. $F(p) = \dfrac{p}{p^3 + 1}$. 18. $F(p) = \dfrac{e^{-p}}{p(p-1)}$.

Solve the Cauchy problem for the following differential equations:

19. $x' + x = e^{-t}$, $x(0) = 1$. **20.** $x'' + x = t$, $x(0) = 0$, $x'(0) = 1$. **21.** $x'' + x = 0$, $x(0) = 1$, $x'(0) = 0$.

22. $x'' + x = 1$, $x(0) = -1$, $x'(0) = 0$. **23.** $x'' + 2x' + x = t$, $x(0) = x'(0) = 0$.

24. $x'' + x = \cos t$, $x(0) = 0$, $x'(0) = 1$. **25.** $x''' + x' = 1$, $x(0) = x'(0) = x''(0) = 0$.

Solve the Cauchy problem for the following systems of differential equations:

26. $\begin{cases} x' + y = 0, \\ y' + x = 0, \end{cases}$ $x(0) = 1$, $y(0) = -1$.

27. $\begin{cases} x' = -y, \\ y' = x + 2y, \end{cases}$ $x(0) = 1$, $y(0) = 0$.

28. $\begin{cases} x' - x - 2y = t, \\ y' - 2x - y = t, \end{cases}$ $x(0) = 2$, $y(0) = 4$.

Solve the integral equations:

29. $\varphi(t) = \sin t + 2\int_0^t \cos(t - \tau)\varphi(\tau)d\tau.$

30. $\varphi(t) = \sin t + \int_0^t (t - \tau)\varphi(\tau)d\tau.$

Answers

1. Yes. 2. Yes. 3. No. 4. $F(p) = \dfrac{2}{p(p^2 + 4)}$.

5. $F(p) = \dfrac{p(p^2 + m^2 + n^2)}{(p^2 + m^2 + n^2)^2 - 4m^2n^2}$. 6. $F(p) = \dfrac{p^2 - 1}{(p^2 + 1)^2}$. 7. $F(p) = \dfrac{2p^2 + 4p + 8}{(p^2 + 4)^2}$.

8. $F(p) = \dfrac{1}{(p - 2)^2 + 1}$. 9. $F(p) = \dfrac{p + 1}{(p + 1)^2 + 4}$. 10. $F(p) = \dfrac{e^{-ap}}{p}(2 - e^{-ap} - e^{-2ap})$.

11. $F(p) = \dfrac{1 - 2e^{-p} + e^{-2p}}{p^2}$. 12. $F(p) = \dfrac{1 + p - e^{-p}}{p^2(e^p - 1)}$. 13. $f(t) = e^{-2t}\sin t$.

14. $f(t) = (1 - t)e^{-t}$. 15. $f(t) = \dfrac{1}{2}t\sin t$. 16. $f(t) = e^t\cos\sqrt{2}\,t + \dfrac{\sqrt{2}}{2}e^t\sin\sqrt{2}\,t$.

17. $f(t) = \frac{1}{3} e^{\frac{t}{2}} \left(\cos \frac{\sqrt{3}}{2} t + \sqrt{3} \sin \frac{\sqrt{3}}{2} t \right) - \frac{1}{3} e^{-t}$. 18. $f(t) = e^{t-1}\eta(t-1) - \eta(t-1)$.

19. $x(t) = (t+1)e^{-t}$. 20. $x(t) = t$. 21. $x(t) = \cos t$. 22. $x(t) = 1 - 2\cos t$.

23. $x(t) = 2e^{-t} + te^{-t} + t - 2$. 24. $x(t) = \frac{t}{2}\sin t + \sin t$. 25. $x(t) = t - \sin t$.

26. $x(t) = e^t,\ y(t) = -e^t$. 27. $x(t) = e^t - te^t,\ y(t) = te^t$.

28. $x(t) = \frac{28}{9} e^{3t} - e^{-t} - \frac{t}{3} - \frac{1}{9},\ y(t) = \frac{28}{9} e^{3t} + e^{-t} - \frac{t}{3} - \frac{1}{9}$. 29. $\varphi(t) = te^t$.

30. $\varphi(t) = \frac{1}{2}(\sinh t + \sin t)$.

Chapter 29

Partial Differential Equations

29.1 Essentials. Examples

We define a partial differential equation to be

$$\mathscr{F}\left(x_1, x_2, \ldots, x_n, u, \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}, \ldots, \frac{\partial^m u}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}}\right) = 0. \tag{29.1}$$

It relates independent variables $x_1, x_2, \ldots, x_n$, the desired function $u = u(x_1, x_2, \ldots, x_n)$ and its partial derivatives (there must be at least one of them).

Here $k_1, k_2, \ldots, k_n$ are nonnegative integers, such that $k_1 + k_2 + \ldots + k_n = m$ and $\mathscr{F}$ is a specified function of its arguments.

The order of a differential equation is the highest order of the partial derivatives present. So, if x and y are independent variables and $u = u(x, y)$ is the required function, then

$$y \frac{\partial u}{\partial x} - x \frac{\partial u}{\partial y} = 0$$

is a differential equation of the first order, and

$$\frac{\partial^2 v}{\partial x^2} - \frac{\partial^2 y}{\partial y^2} = 0, \quad \frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial y^2} = e^u$$

are differential equations of the second order.

For the sake of simplicity use is sometimes made of the following notation:

$$u_x \equiv \frac{\partial u}{\partial x}, \quad u_y \equiv \frac{\partial u}{\partial y}, \quad u_{xx} \equiv \frac{\partial^2 u}{\partial x^2}, \quad u_{xy} \equiv \frac{\partial^2 u}{\partial x \, \partial y}, \ \ldots.$$

Definition. Suppose that we are given a partial differential equatin (29.1) of order m.

A solution of equation (29.1) in some domain D of the independent variables $x_1, x_2, \ldots, x_n$ is any function $u = u(x_1, x_2, \ldots, x_n) \in C^m(D)$*), such that substitution of the function and its derivatives into (29.1) turns the equation into identity in $x_1, x_2, \ldots, x_n$ in D.

Examples. (1) Find the solution $u = u(x, y)$ of the equation

$$\frac{\partial u}{\partial x} = 0. \tag{29.2}$$

◄ Equality (29.2) means that the desired function u is independent of x, but can be any function of y, i.e.,

$$u = \varphi(y). \tag{29.3}$$

The solution (29.3) of equation (29.2) contains thus one arbitrary function. This is the *general solution* of equation (29.2). ►

(2) Find the solution $u = u(x, y)$ of the equation

$$\frac{\partial^2 u}{\partial x\, \partial y} = 0. \tag{29.4}$$

◄ We put $\partial u/\partial y = v$. Then, equation (29.4) will become $\partial v/\partial x = 0$. Its general solution will be an arbitrary function $v = \omega(y)$. Since $v = \partial u/\partial y$, we have the equation $\partial u/\partial y = \omega(y)$. Integrating with respect to y (we take x to be a parameter), we obtain

$$u(x, y) = \int \omega(y)\,dy + g(x),$$

where $g(x)$ is an arbitrary function. Since $\omega(y)$ is an arbitrary function, its integral is an arbitrary function as well; we denote it by $f(y)$. As a result, we obtain the solution to (29.4) in the form

$$u(x, y) = f(y) + g(x). \tag{29.5}$$

Here $f(y)$ and $g(x)$ are arbitrary differentiable functions. ►

Solution (29.5) of the second-order partial differential equation (29.4) now contains two arbitrary functions. It is called the general solution of equation (29.4), since any other solution of the equation can be obtained from (29.5) by suitably choosing the functions f and g.

We see thus that partial differential equations have entire families of solutions. But there exist partial differential equations that have quite narrow, or even empty, sets of solutions.

*) $C^m(D)$ is the collection of functions that are continuous in D together with all derivatives through order m.

So, for instance, the set of real solutions of the equation

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 = 0$$

is the only function $u(x, y) = \text{const}$, and the equation

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + 1 = 0$$

has no real solutions at all.

For the time being we will not be interested in seeking particular solutions. We will make precise later which additional conditions must be specified to obtain a particular solution, i.e., a function satisfying both the differential equation and the additional conditions.

29.2 Linear Partial Differential Equations. Properties of Their Solutions

A partial differential equation is said to be *linear*, if it is linear relative to the required function and all its derivatives that enter into the equation; otherwise, the equation is called *nonlinear*. For example

$$\frac{\partial^2 u}{\partial x^2} = x^2 \frac{\partial^2 u}{\partial y^2} + e^{-x^2}$$

is a linear equation; the equations

$$y \frac{\partial u}{\partial x} - x \frac{\partial u}{\partial y} + u^2 = 0,$$

$$u \frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial y} = x^2 y$$

are nonlinear.

In the general case, a linear differential equation of the second order for a function of two independent variables x, y has the form

$$A(x, y) \frac{\partial^2 u}{\partial x^2} + 2B(x, y) \frac{\partial^2 u}{\partial x \, \partial y} + C(x, y) \frac{\partial^2 u}{\partial y^2} + a(x, y) \frac{\partial u}{\partial x} + b(x, y) \frac{\partial u}{\partial y} + c(x, y)u = f(x, y), \tag{29.6}$$

where $A(x, y)$, $B(x, y)$, $\ldots$, $c(x, y)$, $f(x, y)$ are functions of x, y defined in some domain D on the xy-plane.

If $f(x, y) \equiv 0$ in D, equation (29.6) is said to be homogeneous, otherwise it is inhomogeneous.

Denoting the left-hand side of (29.6) by $L[u]$, we write (29.6) in the form

$$L[u] = f(x, y). \tag{29.7}$$

The corresponding homogeneous equation will be

$$L[u] = 0. \tag{29.8}$$

Here L is a linear differential operator defined always in a linear space $C^2(D)$ by the function $u = u(x, y)$.

Using the linearity property of the operator L, we obtain the following theorems, which express the properties of solutions of linear homogeneous partial differential equations.

Theorem 29.1. *If $u(x, y)$ is a solution of the linear homogeneous equation $L[u] = 0$, then $cu(x, y)$, where c is any constant, is also a solution of the equation.*

Theorem 29.2. *If $u_1(x, y)$ and $u_2(x, y)$ are solutions of the linear homogeneous equation $L[u] = 0$, then the sum $u_1(x, y) + u_2(x, y)$ is also a solution of the equation.*

Corollary. If each of the functions $u_1(x, y)$, $u_2(x, y)$, $\ldots$, $u_k(x, y)$ is a solution of the equation $L[u] = 0$, then the linear combination

$$c_1u_1(x, y) + c_2u_2(x, y) + \ldots + c_ku_k(x, y),$$

where c_1, c_2, $\ldots$, c_k are arbitrary constants, is also a solution of the equation.

These properties are also inherent in solutions of linear homogeneous ordinary differential equations. But a linear homogeneous ordinary differential equation of order n has exactly n linearly independent particular solutions, whose combination gives the general solution of the equation.

Partial differential equations can have an infinite set of linearly independent particular solutions, i.e., a set of solutions such that any finite number of them will be linearly independent functions. Consider a simple example. The equation $\partial u/\partial y = 0$ has the general solution $u = \varphi(x)$, so that its solutions will be, say, functions $1, x, \ldots, x^n, \ldots$. In this connection, in linear problems involving partial differential equations we will have to deal not only with linear combinations of a finite number of solutions, but also with the series $\sum_{n=1}^{\infty} c_nu_n(x, y)$, whose terms are products of some constants c_n and solutions $u_n(x, y)$ of the differential equation.

Cases are possible where a function $u(x, y; \lambda)$ for all values of the parameter λ in some interval (λ_0, λ_1), finite or infinite, is a solution of the equation $L[u] = 0$. We say then that solutions of the equation depend on the constantly varying parameter λ.

If we take a function $C(\lambda)$ such that the first and second derivatives of the integral

$$U(x, y) = \int_{\lambda_0}^{\lambda_1} C(\lambda)u(x, y; \lambda)d\lambda$$

with respect to x and y can be obtained using differentiation under the integration sign, then the integral will be a solution of the equation $L[u) = 0$.

Concerning the linear inhomogeneous equation $L[u] = f$, we can make the following statements:

Theorem 29.3. *If $u(x, y)$ is a solution of the linear inhomogeneous equation $L[u] = f$, and $v(x, y)$ is a solution of the corresponding homogeneous equation $L[u] = 0$, then the sum $u + v$ is a solution of the inhomogeneous equation.*

Theorem 29.4 (on superposition). *If $u_1(x, y)$ is a solution of the equation $L[u] = f_1$ and $u_2(x, y)$ is a solution of the equation $L[u] = f_2$, then $u_1 + u_2$ is a solution of the equation $L[u] = f_1 + f_2$.*

29.3 Classification of Second-Order Linear Differential Equations in Two Independent Variables

Definition. The linear partial differential equation of the second order

$$A(x, y)\frac{\partial^2 u}{\partial x^2} + 2B(x, y)\frac{\partial^2 u}{\partial x\,\partial y} + C(x, y)\frac{\partial^2 u}{\partial y^2} + a(x, y)\frac{\partial u}{\partial x} + b(x, y)\frac{\partial u}{\partial y} + c(x, y)u = f(x, y) \tag{29.9}$$

in some region Ω in the xy-plane is said to be:

(1) hyperbolic in Ω, if

$$\Delta = B^2 - AC > 0 \quad \text{in} \quad \Omega;$$

(2) parabolic in Ω, if

$$\Delta = B^2 - AC \equiv 0 \quad \text{in} \quad \Omega;$$

(3) elliptic in Ω, if

$$\Delta = B^2 - AC < 0 \quad \text{in} \quad \Omega.$$

Using this definition, we can readily verify that the equations $\partial^2 u/\partial x^2 = \partial^2 u/\partial y^2$ and $\partial^2 u/\partial x\,\partial y = 0$ are hyperbolic for all x and y, the equation

$\partial u/\partial x = \partial^2 u/\partial y^2$ are parabolic for all x and y, and the equation $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = 0$ is elliptic for all x and y. The equation $y(\partial^2 u/\partial x^2) + \partial^2 u/\partial y^2 = 0$ is elliptic for $y > 0$, parabolic on the line $y = 0$ and hyperbolic in the half-plane $y < 0$.

It can be shown that under certain constraints on the coefficients of equation (29.9) there exists a nonsingular change of independent variables

$$\begin{aligned} \xi &= \varphi(x, y) \\ \eta &= \psi(x, y) \end{aligned} \quad (\varphi, \psi \in C^2)$$

that transforms equation (29.9) to a simpler, canonic form, peculiar for each type of the equation. If equation (29.9) is hyperbolic ($\Delta > 0$), then it transforms to

$$\frac{\partial^2 u}{\partial \xi \, \partial \eta} = F\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial \eta}\right)$$

or

$$\frac{\partial^2 u}{\partial \xi^2} - \frac{\partial^2 u}{\partial \eta^2} = \Phi\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial n}\right).$$

These are two canonic forms of the *hyperbolic* equation.

If equation (29.9) is parabolic ($\Delta \equiv 0$), then it can be transformed to

$$\frac{\partial^2 u}{\partial \eta^2} = \Phi\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial \eta}\right).$$

This is the canonic form of the *parabolic* equation.

If equation (29.9) is elliptic ($\Delta < 0$), then it can be transformed to

$$\frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \eta^2} = \Phi\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial \eta}\right).$$

This is the canonic form of the *elliptic* equation. Here F and Φ are some functions dependent on the required function u, its first derivatives $\partial u/\partial \xi$, $\partial u/\partial \eta$ and independent variables ξ, η. The form of the functions F and Φ is determined by the original equation (29.9). In some cases the canonic form of an equation enables one to find the general solution of the original equation.

As a rule, the reduction of equation (29.9) to canonic form by a change of independent variables is local in nature, i.e., it is only feasible in some sufficiently small neighbourhood of a point $M_0(x_0, y_0)$ under consideration.

When there are more than two independent variables, equations may also be hyperbolic, parabolic and elliptic. For example, at $n = 4$, the simplest canonic forms of the equations will be

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = 0 \quad \text{(hyperbolic)},$$

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = 0 \quad \text{(parabolic)},$$

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0 \quad \text{(elliptic)}.$$

Here $u = u(x, y, z, t)$.

But if there are more than two independent variables, then in the general case of the linear equation with variable coefficients

$$\sum_{i,\, j=1}^{n} a_{ij}(x_1, x_2, \ldots, x_n) \frac{\partial^2 u}{\partial x_i \, \partial x_j} + \sum_{i=1}^{n} b_i(x_1, x_2, \ldots, x_n) \frac{\partial u}{\partial x_i}$$

$$+ c(x_1, x_2, \ldots, x_n) u = f(x_1, x_2, \ldots, x_n)$$

reduction to canonic form is only possible at a given point $M_0(x_1^0, x_2^0, \ldots, x_n^0)$ but not in an arbitrarily small neighbourhood of the point.

We will confine our discussion to linear differential equations of the second order. Such equations occur in a wide variety of physics problems.

So, oscillatory processes of various nature (vibrations of strings, membranes, acoustic vibrations of gas in tubes, electromagnetic oscillations, and so on) are described by hyperbolic equations. The simplest of such equations is the equation of the vibrations of a string (unidimensional wave equation)

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad u = u(x, t). \tag{29.10}$$

Here x is a spatial coordinate, t is time, $a^2 = T/\varrho$, where T is the tension of the string, ϱ is its linear density.

Thermal conduction and diffusion are described by equations of parabolic type. In the unidimensional case the simplest equation of heat conduction has the form

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad u = u(x, t). \tag{29.11}$$

Here $a^2 = k/c\varrho$, where ϱ is the density of a medium, c is the specific heat capacity, k is the thermal conductivity.

Lastly, steady-state processes, i.e., ones in which the required function is time-independent, are described by elliptic equations, a typical example of which is the Laplace equation

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad u = u(x, y). \tag{29.12}$$

We make sure by direct check that a solution of equation (29.10) is any function $u(x, t)$ of the form

$$u(x, t) = \varphi(x - at) + \psi(x + at),$$

where $\varphi(\xi), \psi(\eta) \in C^2$.

It can be shown that solutions to (29.11) are functions of the form

$$u(x, t; \lambda) = Ae^{-a^2\lambda^2 t} \sin(\lambda x + \alpha),$$

where A, α are arbitrary constants, and λ is a numerical parameter. Integrating, the solution $u(x, t; \lambda) = e^{-a^2\lambda^2 t} \cos \lambda x$ of (29.11) with respect to λ from $-\infty$ to $+\infty$, we will obtain the so-called *fundamental* solution $U(x, t) = \sqrt{\pi/t}\,e^{-x^2/4a^2 t}$ of the heat transfer equation.

Finally, we can easily see that the real-valued functions $P_n(x, y)$ and $Q_n(x, y)$ defined by the relation

$$(x + iy)^n = P_n(x, y) + iQ_n(x, y)$$

are solutions to the Laplace equation (29.12) for $n = 0, 1, 2, \ldots$. This result is a particular case of the general statement that the real and imaginary parts of the analytic function $f(z) = u(x, y) + iv(x, y)$ of a complex variable $z = x + iy$ are each a solution of the Laplace equation (29.12).

The equation (29.12) being linear, the series $\sum_{n=0}^{\infty} \alpha_n P_n(x, y)$ and $\sum_{n=0}^{\infty} \beta_n Q_n(x, y)$ will also be solutions to (29.12), if they converge uniformly, just like the series that result from them by double termwise differentiation with respect to each of the arguments x and y.

As regards the simplest, or canonic, form of equations of hyperbolic, parabolic and elliptic types, we can sketch some ideas about their solutions.

Formulation of basic problems for linear differential equations of the second order. To achieve a complete description of a physical process it is not sufficient to have just the appropriate differential equation. We will

also need to know the initial state of the process (initial conditions) and the conditions on the boundary S of the domain $\Omega \subset R^n$ in which the process occurs (boundary conditions). This stems from the inherent non-uniqueness of the solution of a differential equation.

For example, the general solution of the equation $\partial^2 u/\partial x\, \partial y = 0$ has the form $u(x, y) = f(x) + g(y)$, where f and g are arbitrary differentiable functions. Therefore, to isolate the solution describing a given physical process, we will have to specify additional conditions.

We distinguish three main types of problems involving partial differential equations:

(a) the Cauchy problem for hyperbolic and parabolic equations: given are the initial conditions, the domain Ω coincides with the entire space R^n; there are no boundary conditions;

(b) the boundary problem for elliptic equations: given are the boundary conditions on the boundary S of the domain Ω; the initial conditions are absent;

(c) the mixed problem for hyperbolic and parabolic equations: given are the initial and boundary conditions, $\Omega \neq R^n$.

Exercises

Find the general solution of the equations ($u = u(x, y)$):

1. $\dfrac{\partial^2 u}{\partial y^2} = \dfrac{\partial u}{\partial y}$. 2. $\dfrac{\partial^2 u}{\partial x\, \partial y} = 2y \dfrac{\partial u}{\partial x}$. 3. $\dfrac{\partial^2 u}{\partial y^2} = e^{x+y}$. 4. $\dfrac{\partial^2 u}{\partial y^2} = x + y$.

5. $\dfrac{\partial^2 u}{\partial x\, \partial y} + \dfrac{1}{x} \dfrac{\partial u}{\partial y} = 0$. 6. $\dfrac{\partial^n u}{\partial y^n} = 0$.

7. Assuming $u = u(x, y, z)$, solve the equation

$$\frac{\partial^3 u}{\partial x\, \partial y\, \partial z} = 0.$$

Find the domains where the following equations are hyperbolic, parabolic and elliptic:

8. $\dfrac{\partial^2 u}{\partial x^2} + 2 \dfrac{\partial^2 u}{\partial x\, \partial y} - 3 \dfrac{\partial^2 u}{\partial y^2} = 0.$

9. $\dfrac{\partial^2 u}{\partial x^2} - 2x \dfrac{\partial^2 u}{\partial x\, \partial y} + y \dfrac{\partial^2 u}{\partial y^2} = u + 1.$

Answers

1. $u(x, y) = \Phi_1(x) + \Phi_2(x)e^y$. 2. $u(x, y) = \Phi_1(x)e^{y^2} + \Phi_2(x)$.

3. $u(x, y) = e^{x+y} + \Phi_1(x)y + \Phi_2(x)$. 4. $u(x, y) = \dfrac{xy^2}{2} + \dfrac{y^3}{x} + \Phi_1(x)y + \Phi_2(x)$.

5. $u(x, y) = \dfrac{\Phi_1(y)}{x} + \Phi_2(x)$. 6. $u(x, y) = \Phi_1(x)y^{n-1} + \Phi_2(x)y^{n-2} + \ldots + \Phi_n(x)$.

7. $u(x, y, z) = \Phi_1(x, y) + \Phi_2(x, z) + \Phi_3(y, z)$. 8. The equation is hyperbolic throughout. 9. In the region $x^2 - y > 0$ the equation is hyperbolic; in the region $x^2 - y < 0$ it is elliptic; the curve $y = x^2$ consists of parabolic points.

Chapter 30

Hyperbolic Equations

30.1 Essentials

Hyperbolic equations occur in problems associated with oscillatory processes (vibrations of strings or membranes, electromagnetic oscillations, etc.). One characteristic feature of the processes described by hyperbolic equations is the finite speed of propagation of perturbations.

Consider, for example, the equations of an electromagnetic field. Classically, electromagnetic phenomena are described by Maxwell's equations. In the simplest case of a nonconductive, homogeneous, isotropic medium in the absence of charges and currents, the equations will be

$$\operatorname{curl} \mathbf{H} = \frac{\varepsilon}{c}\frac{\partial \mathbf{E}}{\partial t}, \tag{30.1}$$

$$\operatorname{curl} \mathbf{E} = -\frac{\mu}{c}\frac{\partial \mathbf{H}}{\partial t}, \tag{30.2}$$

$$\operatorname{div} \mathbf{H} = 0, \tag{30.3}$$

$$\operatorname{div} \mathbf{E} = 0. \tag{30.4}$$

Here $\mathbf{E}$ and $\mathbf{H}$ are the strengths of the electric and magnetic fields; ε is the permittivity of the medium, μ is the permeability of the medium (ε, μ = const), c is the velocity of light in vacuum.

Differentiating (30.1) with respect to t, we will obtain

$$\frac{\varepsilon}{c}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \operatorname{curl}\frac{\partial \mathbf{H}}{\partial t}.$$

Substituting $\partial \mathbf{H}/\partial t$ from (30.2) gives

$$\frac{\varepsilon\mu}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = -\operatorname{curl}\operatorname{curl} \mathbf{E}. \tag{30.5}$$

It is well known that

$$\operatorname{curl}\operatorname{curl} \mathbf{E} = [\nabla, [\nabla, \mathbf{E}]] = \nabla(\nabla, \mathbf{E}) - \nabla^2\mathbf{E} = \operatorname{grad}\operatorname{div}\mathbf{E} - \Delta\mathbf{E}.$$

Since div E = 0, from (30.5), we have

$$\frac{\varepsilon\mu}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \Delta \mathbf{E}.$$

Thus, for the vector field E we get the equation

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{c^2}{\varepsilon\mu}\Delta \mathbf{E}.$$

This is one of the most fundamental equations of mathematical physics. It is called the *wave equation*.

We can readily find that the vector field **H** obeys exactly the same equation

$$\frac{\partial^2 \mathbf{H}}{\partial t^2} = \frac{c^2}{\varepsilon\mu}\Delta \mathbf{H}.$$

And so each of the components E_x, E_y, E_z and H_x, H_y, H_z of **E** and **H** in this simple case obey the wave equation

$$\frac{\partial^2 u}{\partial t^2} = a^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right). \tag{30.6}$$

This is a *hyperbolic* equation. Here $a = c/\sqrt{\varepsilon\mu}$ is the speed of propagation of the process, so that electromagnetic processes propagate in a nonconducting medium with the velocity $a = c/\sqrt{\varepsilon\mu}$. Specifically, in a vacuum ($\varepsilon = \mu = 1$) they propagate with the velocity of light c.

If $u = u(x, y, t)$, then equation (30.6) becomes

$$\frac{\partial^2 u}{\partial t^2} = a^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right).$$

When $u = u(x, t)$, we obtain the unidimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2}.$$

We will begin our discussion of hyperbolic equations with the unidimensional wave equation (the equation for vibration of a string) $\partial^2 u/\partial t^2 = a^2\dfrac{\partial^2 u}{\partial x^2}$.

We will define the *string* as an ideally flexible thin thread that is elastic only when it is taut and resists extension.

Suppose that the string vibrates in the xy-plane and that the displacement u is perpendicular at any moment of time t to the x-axis. The vibration process can then be described by one function $u(x, t)$ characterizing the vertical displacement of the string.

We will assume the vibrations to be small and will neglect the quantity $(u_x)^2$ as compared with unity. We can show then that if the linear density of the string is ϱ = const and there are no external forces, the equation of free vibrations of the homogeneous string has the form

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2},$$

where $a^2 = T/\varrho$, T is the tension of the string.

30.2 Solution of the Cauchy Problem (Initial Value Problem) for an Infinite String

Running wave method. D'Alembert solution. We would now like to integrate the equation of free vibrations of the homogeneous string

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}. \tag{30.7}$$

Here $u(x, t)$ is the displacement of a point of the string at a time t relative to the equilibrium position. For each value of t the graph of the function $u = u(x, t)$ gives the shape of the string at t.

We introduce new independent variables ξ, η by

$$\begin{aligned} \xi &= x - at, \\ \eta &= x + at. \end{aligned} \tag{30.8}$$

In terms of ξ, η, equation (30.7) becomes

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta} = 0.$$

Really,

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x} = \frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \eta},$$

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) = \frac{\partial^2 u}{\partial \xi^2}\frac{\partial \xi}{\partial x} + \frac{\partial^2 u}{\partial \xi\, \partial \eta}\frac{\partial \eta}{\partial x}$$

$$+ \frac{\partial^2 u}{\partial \eta\, \partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial^2 u}{\partial \eta^2}\frac{\partial \eta}{\partial x} = \frac{\partial^2 u}{\partial \xi^2} + 2\frac{\partial^2 u}{\partial \xi\, \partial \eta} + \frac{\partial^2 u}{\partial \eta^2},$$

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial t} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial t} = a\left(\frac{\partial u}{\partial \eta} - \frac{\partial u}{\partial \xi}\right),$$

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial t}\left(\frac{\partial u}{\partial t}\right) = a^2\frac{\partial^2 u}{\partial \xi^2} - 2a^2\frac{\partial^2 u}{\partial \xi\, \partial \eta} + a^2\frac{\partial^2 u}{\partial \eta^2}.$$

Substituting the expressions for $\partial^2 u/\partial t^2$ and $\partial^2 u/\partial x^2$ into (30.7), we will arrive at

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta} = 0.$$

This equation is fairly simple to integrate. If we write it in the form

$$\frac{\partial}{\partial \eta}\left(\frac{\partial u}{\partial \xi}\right) = 0,$$

we will have $\partial u/\partial \xi = \omega(\xi)$, where $\omega(\xi)$ is an arbitrary function. Integrating the resultant equation with respect to ξ (here η is treated as a parameter), we will find that

$$u = \int \omega(\xi)\, d\xi + \theta_2(\eta),$$

where $\theta_2(\eta)$ is an arbitrary function of η. Putting $\int \omega(\xi)\, d\xi = \theta_1(\xi)$, we obtain

$$u = \theta_1(\xi) + \theta_2(\eta).$$

Returning to the old variables x and t, we get

$$u(x, t) = \theta_1(x - at) + \theta_2(x + at). \tag{30.9}$$

Direct test shows that the function $u(x, t)$ defined by (30.9), where θ_1 and θ_2 are arbitrary, twice continuously differentiable functions, is the solution of (30.7). This is the general solution of the wave equation (30.7), since any solution of (30.7) can be represented in the form (30.9) with an appropriate choice of θ_1 and θ_2. This solution is called the *D'Alembert solution.*

Each term in (30.9) is also a solution of (30.7). The solution

$$u = \theta_1(x - at) \tag{30.10}$$

has the following physical meaning. At $t = 0$ this solution assumes the form $u = \theta_1(x)$ (Fig. 30.1). Imagine an observer who at $t = 0$ has set off from the point $x = c$ on the x-axis and travels along the positive x-axis with a speed a, so that for his abscissa we have $dx/dt = a$, whence $x = at + c$, i.e., $x - at = c$.

Consequently for him $u = \theta_1(x - at) = \theta_1(c) =$ const. In other words, for the observer the displacement u of the string given by (30.10) will at all times be constant and equal to $\theta_1(c)$. Solution (30.10) is thus a *forward wave,* which propagates along the positive x-axis with the speed a. If we take $\theta_1(\xi)$ to be $\sin \xi$, then we will have a sine wave.

The solution $u = \theta_2(x + at)$ is the *return wave,* which propagates along the negative x-axis with the speed a.

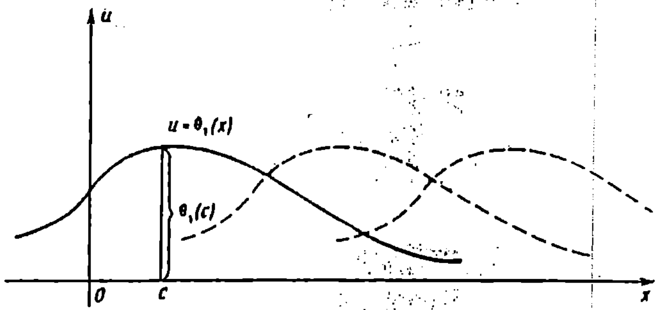


Fig. 30.1

Solution (30.9) is thus the sum of the forward and return waves. This yields the following graphic method of constructing the shape of the string at any moment of time t: we construct the curves $u = \theta_1(x)$ and $u = \theta_2(x)$ representing the forward and return waves at the initial moment of time $t = 0$, and next, without changing their shapes, we shift them simultaneously by $at > 0$ in the opposite directions, the curve $u = \theta_1(x)$ to the right, the curve $u = \theta_2(x)$ to the left.

To obtain the graph of the string it is sufficient to construct the algebraic sum of the ordinates of the shifted curves.

Solution of the Cauchy problem for an infinite string. The Cauchy problem consists in finding the function $u(x, t) \in C^2$, which obeys equation (30.7) for $t > 0$, $-\infty < x < +\infty$, with the initial conditions

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x), \quad -\infty < x < +\infty, \tag{30.11}$$

where $\varphi_0(x) \in C^2(R^1)$, $\varphi_1(x) \in C^1(R^1)$. Here the function $\varphi_0(x)$ defines the shape of the string at $t = 0$; the function $\varphi_1(x)$ gives the distribution of velocities $\partial u/\partial t$ along the string at $t = 0$.

Suppose that the solution of the problem in question exists; it is then given by (30.9).

We define θ_1 and θ_2 so that they obeyed the initial conditions (30.11)

$$u(x, 0) = \theta_1(x) + \theta_2(x) = \varphi_0(x), \tag{30.12}$$

$$u_t(x, 0) = -a[\theta_1'(x) - \theta_2'(x)] = \varphi_1(x). \tag{30.13}$$

Integrating the second of these gives

$$\theta_1(x) - \theta_2(x) = -\frac{1}{a}\int_0^x \varphi_1(\alpha)\,d\alpha + C,$$

where C is an arbitrary constant.

From

$$\theta_1(x) + \theta_2(x) = \varphi_0(x)$$

$$\theta_1(x) - \theta_2(x) = -\frac{1}{a}\int_0^x \varphi_1(\alpha)d\alpha + C,$$

we find

$$\theta_1(x) = \frac{1}{2}\varphi_0(x) - \frac{1}{2a}\int_0^x \varphi_1(\alpha)d\alpha + \frac{C}{2},$$

$$\theta_2(x) = \frac{1}{2}\varphi_0(x) + \frac{1}{2a}\int_0^x \varphi_1(\alpha)d\alpha - \frac{C}{2}.$$

Substituting (30.9) into the expressions for θ_1 and θ_2, we obtain

$$u(x, t) = \frac{1}{2}\varphi_0(x - at) - \frac{1}{2a}\int_0^{x-at} \varphi_1(\alpha)d\alpha$$

$$+ \frac{1}{2}\varphi_0(x + at) + \frac{1}{2a}\int_0^{x+at} \varphi_1(\alpha)d\alpha$$

or

$$u(x, t) = \frac{\varphi_0(x - at) + \varphi_0(x + at)}{2} + \frac{1}{2a}\int_{x-at}^{x+at} \varphi_1(\alpha)d\alpha. \qquad (30.14)$$

This is the so-called *D'Alembert formula.*

It can easily be tested that if $\varphi_0(x) \in C^2(R^1)$, $\varphi_1(x) \in C^1(R^1)$, then the function $u(x, t)$ given by (30.14) obeys (30.7) and the initial conditions (30.11), i.e., it solves the problem.

This solution is unique. Really, if there existed a second solution of the problem (30.7), (30.11), it would be represented by formula (30.14) and coincide with the first solution.

Problem. Using the D'Alembert formula for the solution of the Cauchy problem

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad -\infty < x < +\infty,$$

$$u(x, 0) = \varphi(x), \quad \frac{\partial u(x, 0)}{\partial t} = \psi(x), \quad -\infty < x < +\infty,$$

show that if both $\varphi(x)$ and $\psi(x)$ are odd, $u(x, t)|_{x=0} = 0$, and if they are even, $\partial u/\partial x|_{x=0} = 0$.

Dependence region. It is seen from the D'Alembert formula (30.14) that the value of the solution u at a point P with the coordinates (x, t) is only dependent on the values of φ_0 and φ_1 in the segment γ_p: $[x - at, x + at]$ on the x-axis.

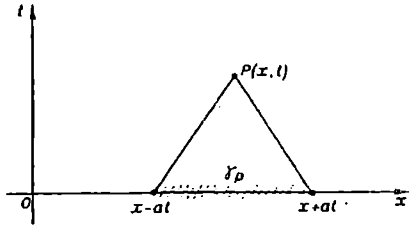


Fig. 30.2

Actually, the solution includes the values of φ_1 on the entire interval γ_p, and the values φ_0 only at the ends of the segment. We say that the solution $u(P)$ "ignores" the conditions of the problem outside γ_p.

The interval γ_p of the x-axis is called the *region of dependence* for the point P (Fig. 30.2).

30.3 Examination of the D'Alembert Formula

Consider two special cases which give some idea of the behaviour of the solution of the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} \tag{30.15}$$

in the general case.

(1) Let $\varphi_1(x) \equiv 0$, and the graph of the function $\varphi_0(x)$ have the form as shown in Fig. 30.3*a*. For simplicity, we will consider that $a = 1$. Then, the D'Alembert formula will be

$$u(x, t) = \frac{\varphi_0(x - t) + \varphi_0(x + t)}{2}.$$

To obtain the graph of $u(x, t)$ viewed as a function of x at some fixed positive t, we proceed as follows: at first we draw two identical coincident graphs, each of which is derived from the plot of $\varphi_0(x)$ by halving the ordinates (the dash line at the top). We then shift one of the plots as a whole

by t to the right along the positive x-axis, and the other by t to the left. We then construct another graph such that the ordinate of each value of x is the sum of the ordinates of the two shifted graphs.

In this way we have constructed the plots of $u(x, 0)$, $u(x, 1/4)$, $u(x, 1/2)$, $u(x, 1)$ in Fig. 30.3.

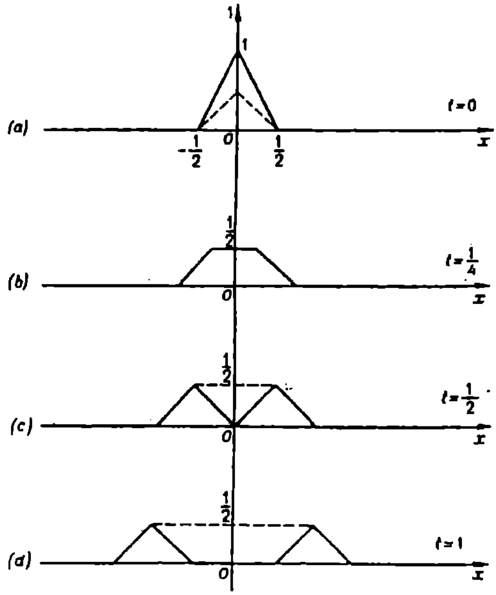


Fig. 30.3

We see that with the initial conditions chosen at each point of the string after both waves have passed (and for points that lie outside of the region of the initial displacement after the passage of a single one) comes a rest.

(2) Let $\varphi_0(x) \equiv 0$, and $\varphi_1(x) = \begin{cases} 1 & \text{for} \quad |x| \leqslant 1/2 \\ 0 & \text{for} \quad |x| > 1/2 \end{cases}$ (Fig. 30.4).

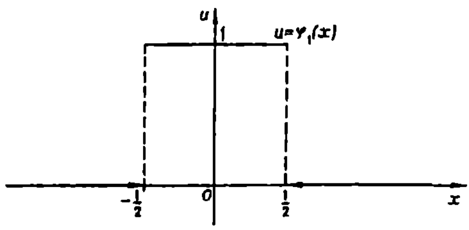


Fig. 30.4

We then say that the string has only the initial pulse. Solution (30.14) becomes (we set $a = 1$)

$$u(x, t) = \frac{1}{2} \int_{x-t}^{x+t} \varphi_1(\alpha)d\alpha.$$

For each fixed x the solution $u(x, t)$ will be zero as long as the intersection of the interval $(x - t, x + t)$ with the interval $(-1/2, 1/2)$, where $\varphi_1(t) \neq 0$ is empty; $u(x, t)$ will vary during the period of time while the growing interval $(x - t, x + t)$ will be embracing ever larger part of the interval $(-1/2, 1/2)$. After the interval $(x - t, x + t)$ will have engulfed the interval $(-1/2, 1/2)$, the quantity $u(x, t)$ will remain unchanged and equal to

$$\frac{1}{2} \int_{-1/2}^{1/2} \varphi_1(\alpha)d\alpha. \tag{30.16}$$

To obtain a curve representing the shape of the string at various t we pro-

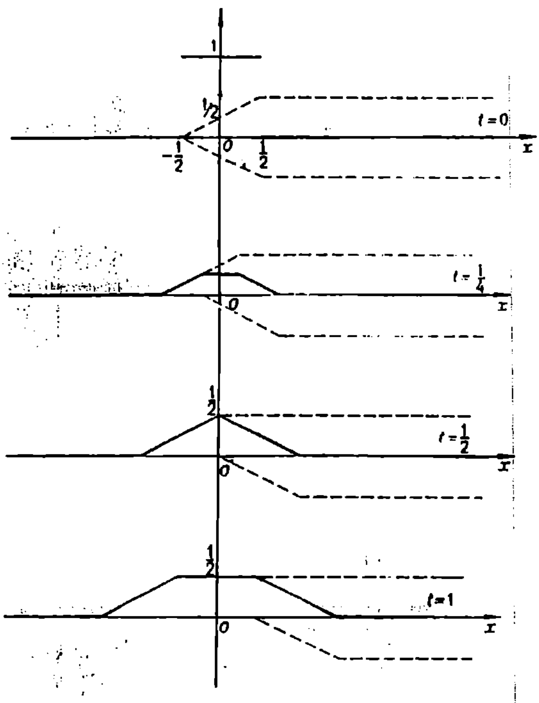


Fig. 30.5

ceed as follows. We denote by $\Phi(z)$ some antiderivative for $\varphi_1(z)$. Then

$$u(x, t) = \frac{1}{2}[\Phi(x+t) - \Phi(x-t)].$$

To plot the curve of $u(x, t)$ we first obtain the graphs of the functions $\frac{1}{2}\Phi(x)$ and $-\frac{1}{2}\Phi(x)$, and then each of these graphs we shift as a whole by t along the x-axis, the first graph to the left, the second to the right. If we add together the ordinates of the shifted graphs, we will obtain the plot of $u(x, t)$ (Fig. 30.5).

After a sufficiently long period of time each point on the string will shift and achieve a stationary displacement u_{st} given by the integral (30.16). We then have a *residual deformation* (*hysteresis*).

30.4 Well-Posedness of a Problem. Hadamard's Example of Ill-Posed Problem

Correctly posed problem. Investigation of physically deterministic phenomena calls for the introduction of the concept of a well-posed problem.

Definition. We say that a mathematical problem is posed *correctly* if
(1) the solution of the problem exists in some class M_1 of functions;
(2) the solution of the problem is unique in some class M_2 of functions;
(3) the solution of the problem is continuously dependent on the conditions of the problem (initial and boundary conditions, coefficients, etc.), i.e., the solution is stable.

The set $M_1 \cap M_2$ of functions is called the *class of correctness*.

In the theory of ordinary differential equations it is proved that the Cauchy problem

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0$$

is well posed, if the function $f(x, y)$ is continuous in the collection of arguments and has a bounded derivative $\partial f/\partial y$ in some domain that contains the point (x_0, y_0).

Consider the Cauchy problem for the infinite string

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad -\infty < x < +\infty, \tag{30.17}$$

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x), \quad -\infty < x < +\infty \tag{30.18}$$

$$\varphi_0(x) \in C^2(R^1), \quad \varphi_1(x) = C^1(R^1).$$

We have established that

(1) the solution of the problem (30.17-18) exists;

(2) the solution is unique.

We show that as the initial conditions vary continuously, so does the solution.

Theorem 30.1. *Whatever the interval* $[0, t_0]$ *of* t *and whatever* $\varepsilon > 0$, *there exists* $\delta = \delta(\varepsilon, t_0) > 0$ *such that for any two solutions* $u_1(x, t)$ *and* $u_2(x, t)$ *of equation* (30.17) *meeting the initial conditions*

$$u_1(x, 0) = \varphi_0(x), \quad u_2(x, 0) = \bar{\varphi}_0(x),$$

$$\frac{\partial u_1(x, 0)}{\partial t} = \varphi_1(x), \quad \frac{\partial u_2(x, 0)}{\partial t} = \bar{\varphi}_1(x),$$

there holds $|u_2(x, t) - u_1(x, t)| < \varepsilon$ *for* $0 \leqslant t \leqslant t_0$, $-\infty < x < +\infty$, *as long as*

$$\begin{cases} |\varphi_0(x) - \bar{\varphi}_0(x)| < \delta \\ |\varphi_1(x) - \bar{\varphi}_1(x)| < \delta \end{cases} \tag{30.19}$$

for $-\infty < x < +\infty$, *i.e., a minor change in the initial conditions causes a minor change in the solutions.*

◄ The functions $u_1(x, t)$ and $u_2(x, t)$ are connected with their initial conditions by the D'Alembert formula, so that

$$u_2(x, t) - u_1(x, t) = \frac{\bar{\varphi}_0(x - at) - \varphi_0(x - at)}{2}$$
$$+ \frac{\bar{\varphi}_0(x + at) - \varphi_0(x + at)}{2} + \frac{1}{2a} \int\limits_{x-at}^{x+at} (\bar{\varphi}_1(\alpha) - \varphi_1(\alpha))\,d\alpha,$$

hence

$$|u_2(x, t) - u_1(x, t)| \leqslant \frac{|\bar{\varphi}_0(x - at) - \varphi_0(x - at)|}{2}$$
$$+ \frac{|\bar{\varphi}_0(x + at) - \varphi_0(x + at)|}{2} + \frac{1}{2a} \int\limits_{x-at}^{x+at} |\bar{\varphi}_1(\alpha) - \varphi_1(\alpha)|\,d\alpha,$$

or, using (30.19),

$$|u_2(x, t) - u_1(x, t)| < \frac{\delta}{2} + \frac{\delta}{2} + \frac{1}{2a}\delta 2at \leqslant \delta(1 + t_0).$$

If we put $\delta = \varepsilon/(1 + t_0)$, we will have from this inequality

$$|u_2(x, t) - u_1(x, t)| < \varepsilon \quad \text{for} \quad 0 \leqslant t \leqslant t_0, \quad -\infty < x < +\infty,$$

which proves the theorem. ►

Therefore, for the wave equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$$

(hyperbolic type) the Cauchy problem is well posed.

Hadamard's example of an ill-posed problem. Consider the following Cauchy problem for the Laplace equation: find the solution of the Laplace equation

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = 0, \quad t > 0, \quad -\infty < x < +\infty \tag{30.20}$$

meeting at $t = 0$ the conditions

$$u|_{t=0} = 0, \tag{30.21}$$

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = \frac{1}{n} \sin nx, \tag{30.22}$$

where n is a natural number.

We can readily check that a solution of the problem will be the function

$$u(x, t) = \frac{1}{n^2} \sinh nt \sin nx. \tag{30.23}$$

Since

$$\left|\frac{\partial u(x, 0)}{\partial t}\right| = \left|\frac{1}{n} \sin nx\right| \leqslant \frac{1}{n},$$

for sufficiently large n the absolute value of $u_t(x, 0)$ will be arbitrarily small throughout. At the same time the solution $u(x, t)$ of the problem under consideration, as follows from (30.23), will have arbitrarily large absolute values for arbitrarily small $t > 0$, if n is sufficiently large.

We assume that we have found the solution $u_0(x, t)$ of the Cauchy problem for equation (30.20) under certain initial conditions

$$u|_{t=0} = \varphi_0(x),$$

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x).$$

Then for the initial conditions

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x) + \frac{1}{n} \sin nx$$

the solution of the Cauchy problem will be

$$u(x, t) = u_0(x, t) + \frac{1}{n^2} \sinh nt \sin nx.$$

It follows that a minor change in the initial conditions may result in arbitrarily large changes in the solution of the Cauchy problem, and in any neighbourhood of the line of initial values $t = 0$.

Therefore, the Cauchy problem for the Laplace equation (elliptic type) is ill posed.

On the other hand, we return to the hyperbolic equation

$$u_{xy} = 0 \tag{30.24}$$

and formulate the following problem: find the solution $u(x, y)$ of the equation (30.24) in the rectangle Q with sides parallel to the coordinate axes (Fig. 30.6), such that on the boundary Γ of that rectangle it assumes the specified boundary values (Γ is closed). This boundary problem, generally speaking, has no solution. We will see this.

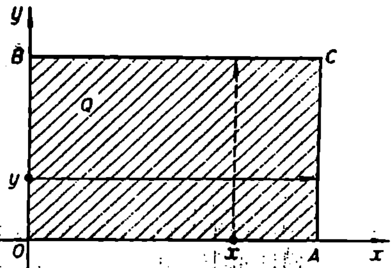


Fig. 30.6

The general solution of (30.24) has the form

$$u(x, y) = f(x) + g(y),$$

where $f, g \in C^1$ are arbitrary functions. Moreover, we cannot arbitrarily specify the boundary values, since the derivative $u_y = g'$ must take on the same values at respective opposite points of the sides $x = \text{const}$. The same is true of the opposite points of the sides $y = \text{const}$.

The values of $u(x, y)$ can only be specified arbitrarily on two adjacent sides of the rectangle (e.g., on OA and OB), but not on the entire boundary Γ of the rectangle, so that for the hyperbolic equation the formulated boundary problem appears to be overdetermined.

In a word, we should not try to make the solution of the above equation obey boundary conditions of an arbitrary type. We should not try to poke a square peg into a round hole, so to speak.

Remark. Ill-posed problems often occur in applications, among them there are many well-known mathematical problems. In particular, the above Cauchy problem for the Laplace equation is concerned with the inverse

problem of gravimetry on determination of the shape of a body from a gravitational anomaly it produces.

30.5 Free Vibrations of a String Fixed at Both Ends. Fourier Method

The Fourier method or the method of separation of variables is one of the commonest methods of solving partial differential equations. Consider it beginning with the simplest problem on free vibrations of a homogeneous string of length l fastened at the ends.

The problem comes down to solving the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad 0 < x < l, \tag{30.25}$$

subject to the boundary conditions

$$u|_{x=0} = u|_{x=l} = 0, \quad t \geqslant 0, \tag{30.26}$$

and the initial conditions

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x), \quad 0 \leqslant x \leqslant l. \tag{30.27}$$

Problem (30.25-27) is called a *mixed problem*, since it contains both initial and boundary conditions.

We will seek particular solutions of equation (30.25) that are not identically zero and meet the boundary conditions (30.26) in the form

$$u(x, t) = T(t)X(x). \tag{30.28}$$

Substituting $u(x, t)$ in the form (30.28) into (30.25) gives

$$T''(t)X(x) = a^2 T(t)X(x),$$

or

$$\frac{T''(t)}{a^2 T(t)} = \frac{X''(x)}{X(x)}.$$

The left-hand side of this equation is only dependent on t, and the right-hand side only on x, and so it is solely possible if its both sides are dependent neither on t nor on x, i.e., both sides are equal to the same constant. We denote this constant by $-\lambda$ (separation constant). We thus have

$$\frac{T''(t)}{a^2 T(x)} = \frac{X''(x)}{X(t)} = -\lambda. \tag{30.29}$$

From this we will then obtain two ordinary differential equations

$$T''(t) + \lambda a^2 T(t) = 0, \tag{30.30}$$

$$X''(x) + \lambda X(x) = 0. \tag{30.31}$$

The boundary conditions (30.26) give

$$u(0, t) = x(0)T(t) = 0,$$
$$u(l, t) = X(l)T(l) = 0.$$

Since $T(t) \not\equiv 0$, we deduce from this that the functions $X(x)$ must meet the boundary conditions

$$X(0) = 0, \quad X(l) = 0. \tag{30.32}$$

Problem (30.31-32) has the obvious solution $X(x) \equiv 0$ (trivial solution). In order to obtain the nontrivial solution $u(x, t)$ of the type (30.28) meeting the boundary conditions (30.26) it is necessary to find the nontrivial solutions of the equation (30.31), satisfying the boundary conditions (30.32).

We thus arrive at the following problem of finding the value of λ at which there exist nontrivial solutions of problem (30.31-32) and also the solutions themselves.

Such values of λ are known as *eigenvalues* (or *proper values*), and the corresponding nontrivial solutions as *eigenfunctions* (or *proper functions*) of the problem.

Formulated that way, the problem is called the *Sturm-Liouville problem*.

Now let us find the eigenvalues and eigenfunctions of problem (30.31-32). We will consider in some detail three cases: $\lambda < 0$, $\lambda = 0$, $\lambda > 0$.

(1) For $\lambda < 0$ the general solution of (30.31) has the form

$$X(x) = C_1 e^{-\sqrt{-\lambda}x} + C_2 e^{\sqrt{-\lambda}x}.$$

Requiring that the boundary conditions (30.32) be met, we will have

$$\begin{cases} C_1 + C_2 = 0, \\ C_1 e^{-\sqrt{-\lambda}l} + C_2 e^{\sqrt{-\lambda}l} = 0. \end{cases} \tag{30.33}$$

Since the determinant of (30.33) is nonzero, then $C_1 = 0$ and $C_2 = 0$. Consequently, $X(x) \equiv 0$, i.e., for $\lambda < 0$ there exist no nontrivial solutions.

(2) For $\lambda = 0$ the general solution of (30.31) has the form

$$X(x) = C_1 x + C_2.$$

The boundary conditions (30.32) yield

$$\begin{cases} C_1 \cdot 0 + C_2 = 0, \\ C_1 \cdot l + C_2 = 0. \end{cases}$$

Hence $C_1 = C_2 = 0$, and so $X(x) \equiv 0$, i.e., at $\lambda = 0$ there are also no nontrivial solutions to the problem.

(3) For $\lambda > 0$ the general solution of equation (30.31) has the form

$$X(x) = C_1 \cos \sqrt{\lambda} x + C_2 \sin \sqrt{\lambda} x.$$

Requiring that the boundary conditions (30.32) be met, we obtain

$$\begin{cases} C_1 \cdot 1 + C_2 \cdot 0 = 0, \\ C_1 \cos \sqrt{\lambda} l + C_2 \sin \sqrt{\lambda} l = 0. \end{cases} \tag{30.34}$$

System (30.34) has nontrivial solutions if and only if the determinant of the system is zero:

$$\begin{vmatrix} 1 & 0 \\ \cos \sqrt{\lambda} l & \sin \sqrt{\lambda} l \end{vmatrix} = 0$$

or $\sin \sqrt{\lambda} l = 0$, hence $\sqrt{\lambda} l = k$, where k is any integer.

The nontrivial solutions of the problem are thus only possible for the values

$$\lambda_k = \left(\frac{k\pi}{l}\right)^2 \quad (k = \pm 1, \ \pm 2, \ \ldots).$$

These are the eigenvalues of the problem (30.31-32).

From the first of the equations (30.34) we find that $C_1 = 0$, and hence the functions

$$X_k(x) = \sin \frac{k\pi}{l} x$$

will be the eigenfunctions of the problem determined up to a factor, which we have taken to be unity.

Corresponding to positive and negative values of k of the same absolute value are eigenfunctions that only differ in a constant factor. Therefore, it is sufficient to take for k only positive integral values: $k = 1, 2, 3, \ldots, n \ldots$.

At $\lambda = \lambda_k$ the general solution of equation (30.30) has the form

$$T_k(t) = A_k \cos \frac{k\pi a}{l} t + B_k \sin \frac{k\pi a}{l} t,$$

where A_k and B_k are arbitrary constants and $A_{k^2} + B_{k^2} > 0$. And so the functions

$$u_k(x, t) = X_k(x) T_k(t) = \left(A_k \cos \frac{k\pi a}{l} t + B_k \sin \frac{k\pi a}{l} t\right) \sin \frac{k\pi}{l} x$$

obey equation (30.25) and boundary conditions (30.26) for any A_k and B_k, $k = 1, 2, \ldots, n, \ldots$.

Equation (30.25) being linear and homogeneous, any finite sum of solutions will also be a solution of the equation. The same holds for the sum of the series

$$u(x, t) = \sum_{k=1}^{\infty} \left(A_k \cos \frac{k\pi a}{l} t + B_k \sin \frac{k\pi a}{l} t\right) \sin \frac{k\pi}{l} x, \tag{30.35}$$

if it converges uniformly and it can be differentiated twice term by term with respect to x and t. Since each term in series (30.35) meets the boundary conditions (30.26), these conditions will also be satisfied by the sum $u(x, t)$ of the series. It now remains to find the constants A_k and B_k in (30.35) so that to meet the initial conditions (30.27).

We differentiate formally the series (30.35) with respect to t

$$\frac{\partial u}{\partial t} = \sum_{k=1}^{\infty} \frac{k\pi a}{l} \left(-A_k \sin \frac{k\pi a}{l} t + B_k \cos \frac{k\pi a}{l} t\right) \sin \frac{k\pi}{l} x. \tag{30.36}$$

Putting in (30.35) and (30.36) $t = 0$, we will, by the initial conditions (30.27), have

$$\begin{aligned} \varphi_0(x) &= \sum_{k=1}^{\infty} A_k \sin \frac{k\pi}{l} x, \\ \varphi_1(x) &= \sum_{k=1}^{\infty} \frac{k\pi a}{l} B_k \sin \frac{k\pi}{l} x. \end{aligned} \tag{30.37}$$

Formulas (30.37) are expansions of the specified functions $\varphi_0(x)$ and $\varphi_1(x)$ into a Fourier series in sines in the interval $(0, l)$.

The coefficients of the expansion (30.37) are worked out from the known formulas (see Chap. 17)

$$\begin{aligned} A_k &= \frac{2}{l} \int_0^l \varphi_0(x) \sin \frac{k\pi}{l} x \, dx, \\ B_k &= \frac{2}{k\pi a} \int_0^l \varphi_1(x) \sin \frac{k\pi}{l} x \, dx, \end{aligned} \qquad (k = 1, 2, \ldots). \tag{30.38}$$

Theorem 30.2 *If $\varphi_0(x) \in C^3$ $[0, l]$ and meets the conditions*

$$\varphi_0(0) = \varphi_0(l) = 0, \quad \varphi_0''(0) = \varphi_0''(l) = 0$$

and $\varphi_1(x) \in C^2\,[0, l]$ *and meets the condition*

$$\varphi_1(0) = \varphi_1(l) = 0,$$

then the sum $u(x, t)$ *of the series* (30.35), *where* A_k *and* B_k *are given by formulas* (30.38), *has in each argument continuous partial derivatives up to the second order in the domain* $\{0 < x < l,\ t > 0\}$, *meets equation* (30.25), *the boundary conditions* (30.26) *and the initial conditions* (30.27), *i.e.,* $u(x, t)$ *is a solution of the problem* (30.25-27).

Example. Find the law of the free vibrations of the homogeneous string of length l fixed at both ends, if at $t = 0$ the string has the shape of the parabola $hx\,(l - x)$, $h > 0 = \text{const}$, the initial speed being zero.

◄ We have the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2\,\frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, \quad t > 0 \tag{30.39}$$

under the boundary conditions

$$u|_{x=0} = 0, \quad u|_{x=l} = 0, \quad t \geqslant 0 \tag{30.40}$$

and the initial conditions

$$u|_{t=0} = hx\,(l - x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = 0, \quad 0 \leqslant x \leqslant l. \tag{30.41}$$

Using the Fourier method, we seek the nontrivial solutions of (30.39) meeting the boundary conditions (30.40) in the form

$$u(x, t) = T(t)\,X(x). \tag{30.42}$$

Substituting $u(x, t)$ in the form (30.42) into (30.39) and separating the variables, we obtain

$$\frac{T''(t)}{a^2 T(t)} = \frac{X''(x)}{X(x)} = -\lambda. \tag{30.43}$$

Hence

$$T''(t) + \lambda a^2 T(t) = 0, \tag{30.44}$$

$$X''(x) + \lambda X(x) = 0. \tag{30.45}$$

From this, by (30.40), we get

$$X(0) = 0, \quad X(l) = 0. \tag{30.46}$$

It has already been established that the eigenvalues for the problem are

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2 \quad (n = 1, 2, \ldots),$$

and the corresponding eigenfunctions are

$$X_n(x) = \sin \frac{n\pi}{l} x \quad (n = 1, 2, \ldots).$$

For $\lambda = \lambda_n$ the general solution of (30.44) is

$$T_n(t) = A_n \cos \frac{n\pi a}{l} t + B_n \sin \frac{n\pi a}{l} t.$$

We will seek the solution of the original problem in the form of the series

$$u(x, t) = \sum_{n=1}^{\infty} \left(A_n \cos \frac{n\pi a}{l} t + B_n \sin \frac{n\pi a}{l} t\right) \sin \frac{n\pi}{l} x. \tag{30.47}$$

To determine the coefficients A_n and B_n we will make use of the initial conditions (30.41)

$$u|_{t=0} = hx(l - x) = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi}{l} x \tag{30.48}$$

$$(0 \leqslant x \leqslant l)$$

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = 0 = \frac{a\pi}{l} \sum_{n=1}^{\infty} n B_n \sin \frac{n\pi}{l} x \tag{30.49}$$

We immediately have from (30.49) that $B_n = 0$ for all n. And from (30.48)

$$A_n = \frac{2h}{l} \int_0^l x (l - x) \sin \frac{n\pi}{l} x \, dx.$$

Integrating twice by parts gives

$$A_{2m+1} = \frac{8l^2 h}{\pi^3 (2m + 1)^3} \quad (m = 0, 1, \ldots).$$

Substituting the found values of A_n and B_n into (30.47), we will obtain the solution of the problem

$$u(x, t) = \frac{8l^2 h}{\pi^3} \sum_{m=0}^{\infty} \frac{1}{(2m + 1)^3} \cos \frac{(2m + 1)\pi a}{l} t$$

$$\times \sin \frac{(2m + 1)\pi}{l} x. \quad ►$$

Remark. If the initial functions $\varphi_0(x)$ and $\varphi_1(x)$ do not meet the conditions of the theorem, then there can be no twice continuously differentiable solution of the mixed problem (30.25-27). But if $\varphi_0(x) \in C^1[0, l]$ and meets the conditions $\varphi_0(0) = \varphi_0(l) = 0$, and $\varphi_1(x) \in C[0, l]$, and $\varphi_1(0) = \varphi_1(l) = 0$, then the series (30.49) converges uniformly for $0 \leqslant x \leqslant l$ and any t and it defines the continuous function $u(x, t)$. In that case, we can speak only about the *generalized* solution of the problem.

Each of the functions

$$u_k(x, t) = T_k(t)X_k(x) = \left(A_k \cos \frac{k\pi a}{l} t + B_k \sin \frac{k\pi a}{l} t\right) \sin \frac{k\pi}{l} x$$

defines the so-called *natural vibrations* of the string fixed at both ends. The string undergoing natural vibrations at $k = 1$ produces the fundamental, i.e., the lowest, tone. For larger k, it produces higher pitches and overtones. Writing $u_k(x, t)$ in the form

$$u_k(x, t) = H_k \sin \frac{k\pi}{l} x \sin \left(\frac{k\pi a}{l} t + \alpha_k\right),$$

we conclude that the natural vibrations of the string are *standing waves* in which points of the string undergo harmonic vibrations with the amplitude $H_k \sin (k\pi x/l)$, frequency $\omega_k = k\pi a/l$ and phase α_k.

We have discussed the case of the free vibrations of the homogeneous string fixed at both ends. Similar results will be obtained if the string is subject to other boundary conditions. Suppose, for instance, that the left end of the string is fixed, i.e., $u(0, t) = 0$, and the right end $x = l$ is elastically linked to its equilibrium position, which corresponds to the condition $u_x(l, t) = -hu(l, t)$ ($h > 0 = \text{const}$). We will again seek the nontrivial solution $u(x, t)$ of the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{30.25}$$

satisfying the established boundary conditions, in the form

$$u(x, t) = T(t)X(x).$$

As a result of the substitution $u(x, t) = T(t)X(x)$ into (30.25) we come to the following eigenvalue problem: find the values of λ for which the differential equation

$$X''(x) + \lambda X(x) = 0 \tag{30.50}$$

with the boundary conditions

$$X(0) = 0, \quad X'(l) + hX(l) = 0 \tag{30.51}$$

has nontrivial solutions $X(x)$.

The general solution of the equation (30.50) has the form ($\lambda > 0$)

$$X(x) = C_1 \cos \sqrt{\lambda} x + C_2 \sin \sqrt{\lambda} x.$$

The first of the boundary conditions (30.51) gives $C_1 = 0$, so that the functions $X(x)$ up to a constant factor are $\sin \sqrt{\lambda} x$.

From the second of (30.51) we have

$$\sqrt{\lambda} \cos \sqrt{\lambda} l + h \sin \sqrt{\lambda} l = 0. \tag{30.52}$$

We put $\lambda = \nu^2$, then, if $h \neq 0$, we have from (30.52),

$$\tan (\nu l) = -\frac{\nu}{h}. \tag{30.53}$$

We thus obtain for ν a transcendental equation. The roots of this equation can be found graphically, by taking the intersections of the successive

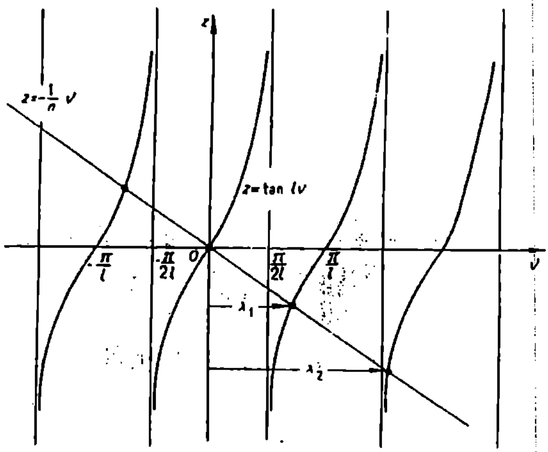


Fig. 30.7

branches of the curve $z = \tan \nu l$ in the (ν, z)-plane by the straight line $z = -\nu/h$ (Fig. 30.7).

Both sides of equation (30.53) are odd functions in ν, therefore corresponding to each positive root ν_k is a negative root of the same absolute value. Since a change of the sign of ν_k does not cause new eigenfunctions to appear (they will only change their sign, which is of no significance), it is sufficient to confine ourselves to the positive roots of equation (30.53). As a result, we again obtain a sequence of eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$, and of the corresponding eigenfunctions $\sin \nu_1 x, \sin \nu_2 x, \ldots, \sin \nu_n x, \ldots$, and also of the natural vibrations $(A_1 \cos a\nu_1 t + B_1 \sin a\nu_1 t) \sin \nu_1 x, \ldots, (A_n \cos a\nu_n t + B_n \sin a\nu_n t) \sin \nu_n x, \ldots$.

By the way, for the nth natural frequency ν_n we have the asymptotic relation

$$\lim_{n\to\infty} \frac{\nu_n}{n} = \frac{\pi}{l}.$$

Specifically, for $l = \pi$, $\lim\limits_{n\to\infty} \nu_n/n = 1$. If the right end of the string $x = l$ is free, i.e., $h = 0$, and hence $u_x(l, t) = 0$, then from (30.52) we obtain $\cos \nu l = 0$. We thus have $\nu l = \pi/2 + n\pi$, so that the eigenvalues will be

$$\nu_n = \frac{(2n+1)\pi}{2l} \quad (n = 0, 1, 2, \ldots),$$

and the corresponding eigenfunctions

$$X_n(x) = \sin \frac{(2n+1)\pi}{2l} x \quad (n = 0, 1, 2, \ldots).$$

30.6 Forced Vibrations of a String Fixed at Both Ends

Consider the vibrations of a homogeneous string of length l fixed at both ends under the action of an external force $f(x, t)$ per unit length. We have the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t) \tag{30.54}$$

with the boundary conditions

$$u|_{x=0} = 0, \quad u|_{x=l} = 0 \tag{30.55}$$

and the initial conditions

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x). \tag{30.56}$$

We will seek the solution $u(x, t)$ of the problem as the sum

$$u(x, t) = v(x, t) + w(x, t), \tag{30.57}$$

where $v(x, t)$ is the solution to the inhomogeneous equation

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2} + f(x, t) \tag{30.58}$$

meeting the boundary conditions

$$v|_{x=0} = 0, \quad v|_{x=l} = 0 \tag{30.59}$$

and the initial conditions

$$v|_{t=0} = 0, \quad \left.\frac{\partial v}{\partial t}\right|_{t=0} = 0 \tag{30.60}$$

and $w(x, t)$ is the solution to the homogeneous equation

$$\frac{\partial^2 w}{\partial t^2} = a^2 \frac{\partial^2 w}{\partial x^2} \tag{30.61}$$

satisfying the boundary conditions

$$w|_{x=0} = 0, \quad w|_{x=l} = 0 \tag{30.62}$$

and the initial conditions

$$w|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial w}{\partial t}\right|_{t=0} = \varphi_1(x). \tag{30.63}$$

The solution $v(x, t)$ represents the *forced* vibrations of the string, i.e., such vibrations which occur under the action of an external disturbing force $f(x, t)$, when the initial perturbations are absent.

Solution $w(x, t)$ represents *free* vibrations of the string, i.e., those which occur only due to initial perturbations.

The method of finding free vibrations $w(x, t)$ has been discussed above, so that it only remains to find the forced vibrations $v(x, t)$, i.e., the solution to the inhomogeneous equation (30.58).

We will apply the method of expanding in eigenfunctions, which is one of the most powerful tools of solving inhomogeneous linear partial differential equations.

The basic idea of the method consists in expanding the external force $f(x, t)$ into a series

$$f(x, t) = \sum_{k=1}^{\infty} f_k(t) X_k(x)$$

in eigenfunctions $\{X_n(x)\}$ of the corresponding homogeneous boundary problem and finding the responses $v_k(x, t)$ of the system to the action of each component $f_k(t)X_k(x)$. Summing up all such responses, we will obtain a solution of the original problem

$$v(x, t) = \sum_{k=1}^{\infty} v_k(x, t).$$

We will seek the solution $v(x, t)$ of the problem (30.58-60) in the form

$$v(x, t) = \sum_{k=1}^{\infty} T_k(t) \sin \frac{k\pi}{l} x. \tag{30.64}$$

Here $\sin k\pi x/l$ are eigenfunctions of the homogeneous boundary problem and the boundary conditions (30.59) are automatically met.

We will define the functions $T_k(t)$ ($k = 1, 2, \ldots$) so that the function $v(x, t)$ would obey equation (30.58) and initial conditions (30.60). Substituting $v(x, t)$ in the form (30.64) into (30.58) gives

$$\sum_{k=1}^{\infty} \left[T_k''(t) + \frac{k^2\pi^2a^2}{l^2} T_k(t) \right] \sin \frac{k\pi}{l} x = f(x, t). \tag{30.65}$$

We expand the function $f(x, t)$ in the interval $(0, l)$ into a Fourier series in sines (eigenfunctions)

$$f(x, t) = \sum_{k=1}^{\infty} f_k(t) \sin \frac{k\pi}{l} x, \tag{30.66}$$

where

$$f_k(t) = \frac{2}{l} \int_0^l f(\xi, t) \sin \frac{k\pi}{l} \xi \, d\xi. \tag{30.67}$$

Comparing expansions (30.65) and (30.66) for the same function $f(x, t)$, we will obtain the differential equations

$$T_k''(t) + \frac{k^2\pi^2a^2}{l^2} T_k(t) = f_k(t) \quad (k = 1, 2, \ldots) \tag{30.68}$$

for the unknown functions $T_k(t)$.

In order that the solution $v(x, t)$ defined by series (30.64) may satisfy the initial conditions (30.60), it is sufficient to require that $T_k(t)$ obey the conditions

$$T_k(0) = 0, \quad T_k'(0) = 0 \quad (k = 1, 2, \ldots). \tag{30.69}$$

We will see that this is so. Putting in (30.64) $t = 0$, we will obtain

$$v(x, 0) = 0 = \sum_{k=1}^{\infty} T_k(0) \sin \frac{k\pi}{l} x \Rightarrow T_k(0) = 0 \quad \forall k.$$

Differentiating (30.64) with respect to t and putting $t = 0$, we will find that $T'(0) = 0$ for all k.

Using, say, the method of variation of constants, we will find that the solutions of equation (30.68) for the initial conditions (30.69) have the form

$$T_k(t) = \frac{l}{k\pi a} \int_0^t f_k(\tau) \sin \frac{k\pi a}{l} (t - \tau) d\tau \quad (k = 1, 2, \ldots), \tag{30.70}$$

where $f_k(t)$ are given by (30.67).

Substituting the found expressions for $T_k(t)$ into series (30.64), we obtain the solution $v(x, t)$ of the problem (30.58-60) if the series (30.64) and the series derived from it by double termwise differentiation with respect to x and t converge uniformly.

It can be shown that such a convergence of the series will be provided, if the function $f(x, t)$ is continuous, has continuous partial derivatives with respect to x up to the second order and all the values of t satisfy the condition $f(0, t) = f(l, t) = 0$.

The solution $u(x, t)$ of the original problem (30.25-27) is representable in the form

$$u(x, t) = \sum_{k=1}^{\infty} \left(A_k \cos \frac{k\pi a}{l} t + B_k \sin \frac{k\pi a}{l} t\right) \sin \frac{k\pi}{l} x + \sum_{k=1}^{\infty} T_k(t) \sin \frac{k\pi}{l} x,$$

where the functions $T_k(t)$ are given by (30.70), and

$$A_k = \frac{2}{l} \int_0^l \varphi_0(x) \sin \frac{k\pi}{l} x \, dx, \quad B_k = \frac{2}{k\pi a} \int_0^l \varphi_1(x) \sin \frac{k\pi}{l} x \, dx \quad (k = 1, 2, \ldots).$$

Example. Solve the mixed problem

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + t \sin x, \quad t > 0, \quad 0 < x < \pi, \tag{30.71}$$

$$u|_{x=0} = 0, \quad u|_{x=\pi} = 0, \quad t \geqslant 0, \tag{30.72}$$

$$u|_{t=0} = 0, \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = 0, \quad 0 \leqslant x \leqslant \pi. \tag{30.73}$$

◄ There are no initial perturbations, so that we have a "pure" problem on forced vibrations of the homogeneous string of length π fixed at both ends.

The system of functions $\{\sin nx\}$ is orthogonal on $[0, \pi]$; it is a system of eigenfunctions of the boundary problem $X''(x) + \lambda X(x) = 0$, $X(0) = X(\pi) = 0$ (here $l = \pi$).

We seek the solution of the problem in the form

$$u(x, t) = \sum_{n=1}^{\infty} T_n(t) \sin nx, \tag{30.74}$$

where $T_n(t)$ are unknown functions. Substituting $u(x, t)$ in the form (30.74) into equation (30.71), we will obtain

$$\sum_{n=1}^{\infty} (T_n''(t) + n^2 T_n(t)) \sin nx = t \sin x.$$

It is easily seen that

$$T_1''(t) + T_1(t) = t \tag{30.75}$$

$$T_n''(t) + n^2 T_n(t) = 0 \quad (n = 2, 3, \ldots). \tag{30.76}$$

Using formula (30.74), we by the initial conditions (30.73) obtain

$$u(x, 0) = 0 = \sum_{n=1}^{\infty} T_n(0) \sin nx,$$

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = 0 = \sum_{n=1}^{\infty} T_n'(0) \sin nx,$$

hence

$$T_n(0) = T_n'(0) = 0 \quad (n = 1, 2, \ldots). \tag{30.77}$$

We thus have for $T_1(t)$

$$T_1''(t) + T_1(t) = t, \tag{30.78}$$

$$T_1(0) = T_1'(0) = 0. \tag{30.79}$$

The general solution of (30.78) is

$$T_1(t) = C_1 \cos t + C_2 \sin t + t.$$

Requiring that the initial conditions (30.79) be met, we find $C_1 = 0$, $C_2 = -1$, so that

$$T_1(t) = t - \sin t.$$

For $n \geqslant 2$ we have

$$T_n''(t) + n^2 T_n(t) = 0,$$

$$T_n(0) = T_n'(0) = 0,$$

hence $T_n \equiv 0$ $(n = 2, 3, \ldots)$.

Using formula (30.74) for the solution $u(x, t)$ of the original problem, we will obtain the expression

$$u(x, t) = (t - \sin t) \sin x.$$

30.7 Forced Vibrations of a String with Unfixed Ends

We now consider the forced vibrations of a homogeneous string of length l under the action of the external force $f(x, t)$ per unit length when the ends of the string are not fixed but move according to a given law. The problem boils down to solving the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad t > 0, \quad 0 < x < l, \tag{30.80}$$

with the boundary conditions

$$u|_{x=0} = \psi_1(t), \quad u|_{x=l} = \psi_2(t), \quad t \geqslant 0 \tag{30.81}$$

and the initial conditions

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x), \quad 0 \leqslant x \leqslant l. \tag{30.82}$$

The Fourier method is not directly applicable to this problem, since the boundary conditions (30.81) are inhomogeneous. But the problem can readily be reduced to a problem with zero (homogeneous) boundary conditions.

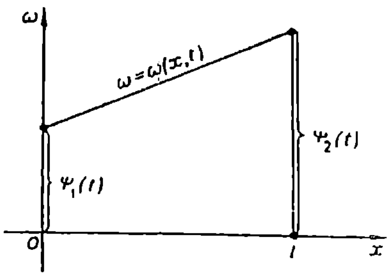


Fig. 30.8

Introduce the auxiliary function

$$\omega(x, t) = \psi_1(t) + [\psi_2(t) - \psi_1(t)]\frac{x}{l}. \tag{30.83}$$

It is easy to see that

$$\omega|_{x=0} = \psi_1(t), \quad \omega|_{x=l} = \psi_2(t). \tag{30.84}$$

Function $\omega(x, t)$ at the ends on the interval $0 \leqslant x \leqslant l$ thus satisfies the conditions (30.81), and inside this interval it is defined linearly in x (Fig. 30.8).

We say that the function $\omega(x, t)$ extends the boundary conditions into the interval $0 < x < l$.

We seek the solution of the problem (30.80-82) as the sum

$$u(x, t) = v(x, t) + \omega(x, t), \tag{30.85}$$

where $v(x, t)$ is a new unknown function.

By virtue of the choice of $\omega(x, t)$, the function $v = u - \omega$ meets the zero boundary conditions

$$v|_{x=0} = (u - \omega)|_{x=0} = 0, \quad v|_{x=l} = (u - \omega)|_{x=l} = 0 \tag{30.86}$$

and the initial conditions

$$\begin{aligned} v|_{t=0} &= u|_{t=0} - \omega|_{t=0} \\ &= \varphi_0(x) - \psi_1(0) - [\psi_2(0) - \psi_1(0)]\frac{x}{l} = \tilde{\varphi}_0(x) \\ \left.\frac{\partial v}{\partial t}\right|_{t=0} &= \left.\frac{\partial u}{\partial t}\right|_{t=0} - \left.\frac{\partial \omega}{\partial t}\right|_{t=0} \\ &= \varphi_1(x) - \psi_1'(0) - [\psi_2'(0) - \psi_1'(0)]\frac{x}{l} = \tilde{\varphi}_1(x). \end{aligned} \tag{30.87}$$

Substituting $u = v + \omega$ into equation (30.80) gives

$$\frac{\partial^2 v}{\partial t^2} = a^2\frac{\partial^2 v}{\partial x^2} + a^2\frac{\partial^2 \omega}{\partial x^2} - \frac{\partial^2 \omega}{\partial t^2} + f(x, t)$$

or, taking into account the expression for $\omega(x, t)$, we obtain

$$\frac{\partial^2 v}{\partial t^2} = a^2\frac{\partial^2 v}{\partial x^2} + f_1(x, t),$$

where

$$f_1(x, t) = f(x, t) - \psi_1''(t) - [\psi_2''(t) - \psi_1''(t)]\frac{x}{l}.$$

If $\psi_1(t), \psi_2(t) \in C^2$, we thus come to the following problem for the function $v(x, t)$: find the solution of the equation

$$\frac{\partial^2 x}{\partial t^2} = a^2\frac{\partial^2 v}{\partial x^2} + f_1(x, t)$$

that meets the boundary conditions

$$v|_{x=0} = 0, \quad v|_{x=l} = 0$$

and the initial conditions

$$v|_{t=0} = \tilde{\varphi}_0(x), \quad \left.\frac{\partial v}{\partial t}\right|_{t=0} = \tilde{\varphi}_1(x),$$

i.e., to a mixed problem with zero boundary conditions. The method of solving such problems has been presented above.

Example. Solve the mixed problem

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad 0 < x < 1, \tag{30.88}$$

$$u|_{x=0} = t, \quad u|_{x=1} = 2t, \quad t \geqslant 0, \tag{30.89}$$

$$u|_{t=0} = 0, \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = 1 + x, \quad 0 \leqslant x \leqslant 1. \tag{30.90}$$

◄ The boundary conditions are inhomogeneous (the ends of the string are not fixed). Here $\psi_1(t) = t$, $\psi_2(t) = 2t$. We introduce the auxiliary function

$$\omega(x, t) = t + tx = t(1 + x). \tag{30.91}$$

We will seek the solution of the problem in the form

$$u(x, t) = v(x, t) + \omega(x, t), \tag{30.92}$$

where $v(x, t)$ is a new unknown function.

For it we will obtain the equation

$$\frac{\partial^2 v}{\partial t^2} = \frac{\partial^2 v}{\partial x^2}, \quad t > 0, \quad 0 < x < 1, \tag{30.93}$$

with the boundary conditions

$$v|_{x=0} = 0, \quad v|_{x=1} = 0 \tag{30.94}$$

and the initial conditions

$$v|_{t=0} = 0, \quad \left.\frac{\partial v}{\partial t}\right|_{t=0} = 0. \tag{30.95}$$

The problem (30.93-95) has the obvious solution $v(x, t) \equiv 0$ and, as follows from physical considerations, this solution is the only one. Then, from (30.92), we will obtain the solution $u(x, t)$ for the original problem:

$$u(x, t) = t(1 + x). \text{ ►}$$

30.8 General Scheme of the Fourier Method

Consider in the domain Q $\{t > 0,\ 0 < x < l\}$ the differential equation

$$\varrho(x)\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(p(x)\frac{\partial u}{\partial x}\right) - q(x)u \tag{30.96}$$

(the equation of vibrations of an inhomogeneous string of length l), where $\varrho(x) > 0$, $p(x) > 0$, $q(x) \geqslant 0$ for $0 \leqslant x \leqslant l$, so that equation (30.96) is hyperbolic in the domain Q. Further, we assume that

$$\varrho(x) \in C[0, l], \quad p(x) \in C^1[0, l], \quad q(x) \in C[0, l].$$

We now turn to the mixed problem for equation (30.96) with the homogeneous boundary conditions

$$\alpha u(0, t) + \beta u_x(0, t) = 0, \quad \gamma u(l, t) + \delta u_x(l, t) = 0, \quad t \geqslant 0,$$

where α, β, γ, δ are some constants, such that $\alpha^2 + \beta^2 \neq 0$, $\gamma^2 + \delta^2 \neq 0$. (Recall that a problem is called homogeneous, if along with the solution u of the problem, cu, where c is an arbitrary constant, is also a solution.)

In particular, boundary conditions of the following types are possible:

(1) $u(0, t) = 0$ and $u(l, t) = 0$ (a string with fixed ends; Fig. 30.9*a*);

(2) $u_x(0, t) = 0$ and $u_x(l, t) = 0$ (a string with free ends; Fig. 30.9*b*);

(3) $u_x(0, t) = h_0 u(0, t)$ and $u_x(l, t) = -h_1 u(l, t)$ (a string with elastically fixed ends; Fig. 30.9*c*).

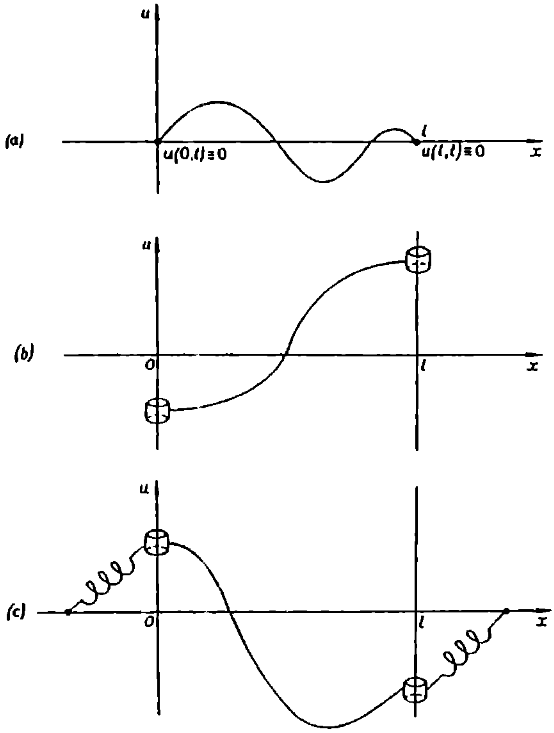


Fig. 30.9

Numbers h_0 and h_1 must be positive, if the position of rest is a stable equilibrium.

For simplicity we will only confine ourselves to the case of the string with fixed ends, and so we come to the following problem: find the solution $u(x, t)$ of the equation

$$\varrho(x)\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(p(x)\frac{\partial u}{\partial x}\right) - q(x)u, \quad t > 0, \quad 0 < x < l, \tag{30.97}$$

obeying the boundary conditions

$$u|_{x=0} = 0, \quad u|_{x=l} = 0, \quad t \geqslant 0 \tag{30.98}$$

and the initial conditions

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x), \quad 0 \leqslant x \leqslant l. \tag{30.99}$$

We will solve the problem by the Fourier method.

(1) We look for nontrivial solutions of equation (30.97) that meet the boundary conditions (30.98) as the product

$$u(x, t) = T(t)X(x). \tag{30.100}$$

Substituting $u(x, t)$ in the form (30.100) into (30.97) gives

$$T(t)\frac{d}{dx}\left[p(x)\frac{dX}{dx}\right] - q(x)X(x)T(t) = \varrho(x)X(x)T''(t)$$

or

$$\frac{\dfrac{d}{dx}\left[p(x)\dfrac{dX}{dx}\right] - q(x)X(x)}{\varrho(x)X(x)} = \frac{T''(t)}{T(t)}. \tag{30.101}$$

The left-hand side of (30.101) is only dependent on x, and the right-hand side only on t, and the equality is only possible when the ratio (30.101) has a constant value. We denote this constant by $-\lambda$. Then, from (30.101), we obtain two ordinary differential equations

$$T''(t) + \lambda T(t) = 0, \tag{30.102}$$

$$\frac{d}{dx}\left[p(x)\frac{dX}{dx}\right] + (\lambda\varrho(x) - q(x))X(x) = 0. \tag{30.103}$$

If we wish to obtain nontrivial solutions of equation (30.97) of the form (30.100) satisfying the boundary conditions (30.98), it is necessary that the function $X(x)$ be a nontrivial solution of (30.103) satisfying the boundary conditions

$$X(0) = 0, \quad X(l) = 0. \tag{30.104}$$

As we have seen, not for any λ the problem has a solution that is not identically zero. Just as in the special case of the homogeneous string, we come to the *Sturm-Liouville problem on eigenvalues*: find the values of λ for which there exist nontrivial solutions of equation (30.103) satisfying the boundary conditions (30.104), as well as the solutions themselves.

The values of λ for which the problem has a nontrivial solution are called the *eigenvalues*, and the solutions themselves the *eigenfunctions* corresponding to a given eigenvalue. The collection of all eigenvalues is called the *spectrum* of a given problem.

Equation (30.103) and boundary conditions (30.104) being homogeneous, the eigenfunctions are determined up to a constant factor. We choose the factor so that

$$\int_0^l \varrho(x)X_k^2(x)dx = 1. \tag{30.105}$$

The eigenfunctions satisfying the condition (30.105) are said to be *normalized with weight* $\varrho(x)$.

We now establish some general properties of the eigenvalues and eigenfunctions of the Sturm-Liouville problem.

Theorem 30.3. *Corresponding to each eigenvalue up to a constant factor is only one eigenfunction* (we exclude problems with periodicity conditions).

◄ Suppose that there exist two eigenfunctions $X_1(x)$ and $X_2(x)$ that correspond to the same eigenvalue λ_0, i.e., ones that obey the differential equation (30.103) for the same $\lambda = \lambda_0$. As stated, $X_1(0) = 0$ and $X_2(0) = 0$, and the Wronskian

$$W(x) = \begin{vmatrix} X_1(x) & X_2(x) \\ X_1'(x) & X_2'(x) \end{vmatrix}$$

for the solutions $X_1(x)$ and $X_2(x)$ of equation (30.103) at the point $x = 0$ becomes zero, and hence the solutions $X_1(x)$ and $X_2(x)$ are linearly dependent. ►

Theorem 30.4. *Eigenfunctions corresponding to various eigenvalues are orthogonal with weight* $\varrho(x)$ *on the interval* $[0, l]$, *i.e.,*

$$\int_0^l \varrho(x)X_m(x)X_n(x)dx = 0,$$

where $X_m(x)$ *and* $X_n(x)$ *are eigenfunctions corresponding to various eigenvalues* λ_m *and* λ_n, $m \neq n$.

◄ To begin with, we establish one proposition that is of interest in its own right. We take the so-called *Sturm-Liouville operator*

$$Ly \equiv \frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) - q(x)y \tag{30.106}$$

where $p(x)$, $p'(x)$, $q(x) \in C[0, l]$, $p(x) \geqslant 0$, $q(x) \geqslant 0$ on $[0, l]$, and consider it on the set $\overset{\circ}{C}{}^2[0, l]$ of functions that are twice continuously differentiable on $[0, l]$ and satisfy the boundary conditions $y(0) = y(l) = 0$

$$\overset{\circ}{C}{}^2[0, l] \doteq \{y(x) | y(x) \in C^2[0, l];\ y(0) = y(l) = 0\}.$$

Lemma. The Sturm-Liouville operator (30.106) on $\overset{\circ}{C}{}^2[0, l]$ is symmetrical

$$(Lu, v) = (u, Lv).$$

Here $u(x)$, $v(x) \in \overset{\circ}{C}{}^2[0, l]$; $(f, g) = \int_0^l f(x)g(x)dx$.

Indeed,

$$(Lu, v) = \int_0^l \left\{\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) - q(x)u\right\} v(x)dx$$

$$= -\int_0^l q(x)v(x)u(x)dx + \int_0^l \frac{d}{dx}\left(p(x)\frac{du}{dx}\right) v\, dx.$$

Integrating by parts the last integral on the right-hand side and considering that $v|_{x=0} = v|_{x=l} = 0$, we find that

$$(Lu, v) = -\int_0^l q(x)u \cdot v\, dx - \int_0^l p(x)\frac{du}{dx}\frac{dv}{dx}dx.$$

Integrating by parts the second term on the right-hand side and considering that $u|_{x=0} = u|_{x=l} = 0$, we get

$$(Lu, v) = -\int_0^l q(x)u \cdot v\, dx + \int_0^l \frac{d}{dx}\left(p(x)\frac{dv}{dx}\right) u\, dx$$

$$= \int_0^l u\left\{\frac{d}{dx}\left(p(x)\frac{dv}{dx}\right) - q(x)v\right\} dx = (u, Lv). \blacktriangleright$$

We now turn to the proof of the theorem. We write equation (30.103) in the form

$$\frac{d}{dx}\left(p(x)\frac{dX}{dx}\right) - q(x)X(x) = -\lambda\varrho(x)X(x) \tag{30.107}$$

and denote by $L[X]$ the operator on the left of (30.107). This is the Sturm-

Liouville operator. On the set of eigenfunctions $X_k(x)$ of the problem (30.103-104) this is a symmetrical operator.

Let $X_m(x)$ be an eigenfunction of the problem (30.103-104) that corresponds to the eigenvalue λ_m, and $X_n(x)$ be an eigenfunction corresponding to the eigenvalue λ_n ($\lambda_n \neq \lambda_m$). We then have

$$\begin{aligned} L[X_m(x)] &\equiv -\lambda_m \varrho(x) X_m(x) \\ L[X_n(x)] &\equiv -\lambda_n \varrho(x) X_n(x) \end{aligned} \qquad (0 < x < l).$$

We multiply the first identity by $X_n(x)$, the second by $X_m(x)$ and integrate the results with respect to x from 0 to l. We obtain

$$(L[X_m], X_n) = -\lambda_m(\varrho X_m, X_n), \tag{30.108}$$

$$(L[X_n], X_m) = -\lambda_n(\varrho X_n, X_m). \tag{30.109}$$

We note that

$$(L[X_m], X_n) = (X_m, L[X_n]) = (L[X_n], X_m)$$

and subtract (30.109) from (30.108) term by term to obtain

$$0 = (\lambda_n - \lambda_m)(\varrho X_m, X_n).$$

Since $\lambda_n \neq \lambda_m$, it follows from this that $(\varrho X_m, X_n) = 0$, or equivalently

$$\int_0^l \varrho(x) X_m(x) X_n(x) dx = 0 \quad (\lambda_n \neq \lambda_m). \blacktriangleright$$

So, in the special case of the homogeneous string ($\varrho = p = 1$, $q = 0$) fixed at both ends, the eigenfunctions $X_n(x) = \sin n\pi x/l$ ($n = 1, 2, \ldots$) form an orthogonal system of functions in the interval $[0, l]$

$$\int_0^l \sin \frac{n\pi}{l} x \sin \frac{m\pi}{l} x \, dx = 0, \quad m \neq n.$$

Theorem 30.5. *All the eigenvalues of problem* (30.103-104) *are real.*

◀ Suppose that there exists a complex eigenvalue $\lambda = \alpha + i\beta$, $\beta \neq 0$, and the corresponding eigenfunction is $X(x) = u(x) + iv(x)$. The complex conjugate number $\bar{\lambda} = \alpha - i\beta$ will then also be an eigenvalue, and the function $\overline{X}(x)$, which is a complex conjugate of $X(x)$, will be the corresponding eigenfunction, since the coefficients of equation (30.103) and the boundary conditions (30.104) are real. Since the eigenfunctions corresponding to various eigenvalues are orthogonal, we have

$$\int_0^l \varrho(x) X(x) \overline{X}(x) dx = \int_0^l \varrho(x) |X(x)|^2 dx = 0,$$

hence $X(x) \equiv 0$, i.e., the complex number λ is not an eigenvalue. ▶

Theorem 30.6. *If $p(x) > 0$, $\varrho(x) > 0$, $q(x) \geqslant 0$ on the interval $[0, l]$, then all the eigenvalues of the problem* (30.103-104) *are positive.*

◄ Let λ_k be an eigenvalue, and $X_k(x)$ be a corresponding eigenfunction normalized with weight $\varrho(x)$. We then have

$$\frac{d}{dx}\left(p(x)\frac{dX_k}{dx}\right) - q(x)X_k(x) \equiv -\lambda_k \varrho(x)X_k(x).$$

Multiplying both sides of this by $X_k(x)$, integrating the result with respect to x from 0 to l and considering that

$$\int_0^l \varrho(x)X_k^2(x)dx = 1$$

we will obtain

$$\lambda_k = \int_0^l q(x)X_k^2(x)dx - \int_0^l \frac{d}{dx}\left(p(x)\frac{dX_k}{dx}\right)X_k(x)dx.$$

Integrating by parts the second term on the right gives

$$\lambda_k = \int_0^l q(x)X_k^2(x)dx + \int_0^l p(x)\left(\frac{dX_k}{dx}\right)^2 dx. \tag{30.110}$$

We find that $dX_k/dt \not\equiv 0$, since otherwise $X_k(x) \equiv \text{const}$ and from the boundary conditions (30.104) we would have $X_k(x) \equiv 0$, which is impossible. The right-hand side of (30.110) is thus positive, which suggests that all the eigenvalues λ_k of the problem are positive. ►

Theorem 30.7. *The problem* (30.103-104) *has a countable set of eigenvalues*

$$\lambda_1 < \lambda_2 < \ldots < \lambda_n < \ldots, \quad \lim_{n\to\infty} \lambda_n = +\infty,$$

with the corresponding eigenfunctions

$$X_1(x),\ X_2(x),\ \ldots,\ X_n(x)\ \ldots.$$

We will carry on our discussion of the Fourier method.

(2) We turn to the differential equation (30.102). Its general solution at $\lambda = \lambda_k$ ($\lambda_k > 0$, see Theorem 30.6) has the form

$$T_k(t) = A_k \cos\sqrt{\lambda_k}\, t + B_k \sin\sqrt{\lambda_k}\, t,$$

where A_k, B_k are arbitrary constants.

Each function

$$u_k(x, t) = T_k(t)X_k(x) = (A_k \cos\sqrt{\lambda_k}\, t + B_k \sin\sqrt{\lambda_k}\, t)X_k(x)$$

will be a solution to (30.97) satisfying the boundary conditions (30.98).

(3) We set up the formal series

$$u(x, t) = \sum_{k=1}^{\infty} (A_k \cos \sqrt{\lambda_k}\, t + B_k \sin \sqrt{\lambda_k}\, t) X_k(x). \tag{30.111}$$

If the series converges uniformly, just like the series obtained from it by double termwise differentiation with respect to x and t, its sum $u(x, t)$ will be a solution of equation (30.97) meeting the boundary conditions (30.98).

To satisfy the initial conditions (30.99) it is needed that

$$u|_{t=0} = \varphi_0(x) = \sum_{k=1}^{\infty} A_k X_k(x), \tag{30.112}$$

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x) = \sum_{k=1}^{\infty} B_k \sqrt{\lambda_k}\, X_k(x). \tag{30.113}$$

We have thus come to the problem on expanding an arbitrary function into a Fourier series in the eigenfunctions $X_k(x)$ of the boundary problem (30.103-104). Assuming that series (30.112) and (30.113) converge uniformly, we can find the coefficients A_k and B_k by multiplying both sides of (30.112) and (30.113) by $\varrho(x)X_n(x)$ and integrating with respect to x from 0 to l. Assuming the functions $X_k(x)$ to be orthonormal with weight $\varrho(x)$ in the interval $[0, l]$, we obtain the following expressions for the Fourier coefficients of the functions $\varphi_0(x)$ and $\varphi_1(x)$ in the system $\{X_k(x)\}$

$$\left.\begin{aligned} A_n &= \int_0^l \varrho(x)\,\varphi_0(x)\,X_n(x)\,dx \\ B_n &= \frac{1}{\sqrt{\lambda_n}} \int_0^l \varrho(x)\,\varphi_1(x)\,X_n(x)\,dx \end{aligned}\right\} \quad (n = 1, 2, \ldots).$$

We will look for A_n and B_n leaning on the Steklov theorem on expansion.

Theorem 30.8. *Any twice continuously differentiable function $F(x)$ meeting the boundary conditions of the problem can be expanded into an absolutely and uniformly convergent series in the eigenfunctions $X_k(x)$ of the problem:*

$$F(x) = \sum_{n=1}^{\infty} c_n X_n(x),$$

where

$$c_n = \int_0^l \varrho(x)\,F(x)\,X_n(x)\,dx,$$

and $X_n(x)$ $(n = 1, 2, \ldots)$ are the eigenfunctions normalized with weight $\varrho(x)$.

Substituting the found values of A_n and B_n into series (30.111), we obtain the solution $u(x, t)$ of the mixed problem (30.97-99), if series (30.111) and the series obtained from it by double termwise differentiation with respect to x and t converge uniformly.

Remark. We have considered the case of the simplest boundary conditions $u(0, t) = u(l, t) = 0$. By slightly changing the above treatment, we can prove the analogous properties of the eigenvalues and eigenfunctions of the more general homogeneous boundary problem

$$\alpha u(0, t) + \beta u_x(0, t) = 0, \quad \gamma u(l, t) + \delta u_x(l, t) = 0$$
$$\alpha^2 + \beta^2 \neq 0, \quad \gamma^2 + \delta^2 \neq 0.$$

30.9 Uniqueness of Solution of a Mixed Problem

We prove the uniqueness of the solution of the mixed problem for the forced vibration of a homogeneous string

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad t > 0, \quad 0 < x < l, \tag{30.114}$$

$$u(0, t) = \psi_1(t), \quad u(l, t) = \psi_2(t), \quad t \geqslant 0, \tag{30.115}$$

$$u|_{t=0} = \varphi_0(x), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = \varphi_1(x), \quad 0 \leqslant x \leqslant l. \tag{30.116}$$

Suppose that there exist two solutions $u_1(x, t)$ and $u_2(x, t)$ of the problem (30.114-117). Then the difference $v(x, t) = u_1(x, t) - u_2(x, t)$ of these solutions will obey the homogeneous equation

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2}, \quad t > 0, \quad 0 < x < l, \tag{30.117}$$

with the zero boundary conditions

$$v(0, t) = 0, \quad v(l, t) = 0, \quad t \geqslant 0 \tag{30.118}$$

and zero initial conditions

$$v(x, 0) = 0, \quad v_t(x, 0) = 0, \quad 0 \leqslant x \leqslant l. \tag{30.119}$$

We will now show that the relations (30.117-119) are only obeyed by the function that is identically zero.

Consider the function

$$E(t) = \frac{1}{2} \int_0^l \left[\left(\frac{\partial v}{\partial t} \right)^2 + a^2 \left(\frac{\partial v}{\partial x} \right)^2 \right] dx \tag{30.120}$$

and show that, provided it meets the conditions (30.117-119), it is independent of t. Differentiating with respect to t gives

$$\frac{dE(t)}{dt} = \int_0^l \left[\frac{\partial v}{\partial t}\frac{\partial^2 v}{\partial t^2} + a^2 \frac{\partial v}{\partial x}\frac{\partial^2 v}{\partial x\, \partial t}\right] dx$$

$$= \int_0^l \frac{\partial v}{\partial t}\frac{\partial^2 v}{\partial t^2}\, dx + \left(a^2 \frac{\partial v}{\partial x}\frac{\partial v}{\partial t}\right)\Bigg|_{x=0}^{x=l} - a^2 \int_0^l \frac{\partial v}{\partial t}\frac{\partial^2 v}{\partial x^2}\, dx$$

$$= \int_0^l \frac{\partial v}{\partial t}\left[\frac{\partial^2 v}{\partial t^2} - a^2 \frac{\partial^2 v}{\partial t^2}\right] dx \equiv 0,$$

since the second term vanishes by virtue of the conditions (30.118) and $\dfrac{\partial^2 v}{\partial t^2} - a^2 \dfrac{\partial^2 v}{\partial x^2} \equiv 0$, since $v(x, t)$ is a solution of equation (30.117).

Therefore, $\dfrac{dE(t)}{dt} \equiv 0$, i.e., $E(t) \equiv \text{const}$. Taking into account the initial conditions (30.119), we will have

$$E(0) = \frac{1}{2}\int_0^l \left[\left(\frac{\partial v}{\partial t}\right)^2 + a^2 \left(\frac{\partial v}{\partial x}\right)^2\right]\Bigg|_{t=0} dx = 0,$$

and so $E(t) \equiv 0$. The integral of the continuous nonnegative function is zero:

$$\int_0^l \left[\left(\frac{\partial v}{\partial t}\right)^2 + a^2 \left(\frac{\partial v}{\partial x}\right)^2\right] dx = 0$$

and so

$$\left(\frac{\partial v}{\partial t}\right)^2 + a^2 \left(\frac{\partial v}{\partial x}\right)^2 \equiv 0.$$

It follows that $\partial v/\partial t \equiv 0$ and $\partial v/\partial x \equiv 0$, so that $v(x, t) \equiv \text{const}$. By virtue of the first of the initial conditions (30.119), $v(x, 0) = 0$, and hence $v(x, t) \equiv 0$, i.e., $u_1(x, t) \equiv u_2(x, t)$.

The integral (30.120) can be rewritten as ($a^2 = T/\varrho$)

$$\frac{1}{2}\int_0^l \left[\left(\frac{\partial v}{\partial t}\right)^2 + a^2 \left(\frac{\partial v}{\partial x}\right)^2\right] dx = \frac{1}{\varrho}\int_0^l \left(\frac{1}{2}\varrho v_t^2 + \frac{1}{2} T v_x^2\right) dx.$$

The quantity $\int_0^l \frac{1}{2} \varrho v_t^2 \, dx$ is the kinetic energy of the string at a time t, and $\int_0^l \frac{1}{2} T v_x^2 \, dx$ is its potential energy, so that the function $E(t)$ up to the factor $\varrho^{-1} = \text{const}$ is the total energy of the string. The equality $E(t) \equiv 0$ is the mathematical expression of the energy conservation law for free oscillations of any physical nature under zero boundary conditions, i.e., when there is no input or dissipation of energy in the process of oscillations. The inhomogeneity of boundary conditions and the inhomogeneity of the equation attest to the presence of constantly active factors feeding or dissipating energy. The inhomogeneity of the initial conditions implies that at the initial moment of time the process has some store of energy, which it retains throughout the oscillation.

The above treatment of the uniqueness of the solution of the mixed problem is said to be energy-based and it is widely employed to establish various uniqueness theorems.

30.10 Vibrations of a Round Membrane

The Fourier method of separation of variables is also used to study the oscillations of bounded bodies, plane or having a volume. By way of example, we consider the problem on the free vibrations of a homogeneous round membrane of radius r_0 with centre at the origin of coordinates, fixed along the edge, that are caused by initial perturbations.

The equation of the vibrations will then be

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right).$$

We introduce the polar coordinates r and φ. The displacement of the points of the membrane will then be a function of the polar coordinates r and φ, and time t: $u = u(r, \varphi, t)$. The expression for the Laplace operator $\Delta u \equiv \partial^2 u/\partial x^2 + \partial^2 u/\partial y^2$ in the polar coordinates is

$$\Delta u \equiv \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2}$$

and the equation of the membrane vibrations will be written as

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} \right).$$

Therefore, the problem of membrane vibrations is posed as follows: find

the function $u(r, \varphi, t)$ obeying the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2}\right) \tag{30.121}$$

with the boundary conditions

$$u|_{r=r_0} = 0 \tag{30.122}$$

(the membrane is fastened along the edge), and with the initial conditions

$$u|_{t=0} = f(r, \varphi), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = F(r, \varphi). \tag{30.123}$$

We will concentrate on the important particular case of the axisymmetrical vibrations, when the initial functions f and F are independent of φ. It is clear then that at any moment of time $t > 0$ the magnitude of the membrane's displacement will not depend on the polar angle φ and will only be the function of r and t: $u = u(r, t)$. This means that for any fixed t the shape of the vibrating membrane will be a surface of revolution.

Under these assumptions the problem reduces to seeking the solution $u(r, t)$ of the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right) \tag{30.124}$$

under the boundary conditions

$$u|_{r=r_0} = 0 \tag{30.125}$$

and the initial conditions

$$u|_{t=0} = f(r), \quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = F(r). \tag{30.126}$$

Using the method of separation of variables, we will seek the nontrivial solutions of (30.124) satisfying the boundary condition (30.125) in the form

$$u(r, t) = T(t)R(r). \tag{30.127}$$

Substituting the function $u(r, t)$ in the form (30.127) into (30.124) and separating the variables, we will have

$$\frac{T''(t)}{a^2 T(t)} = \frac{R''(r) + \frac{1}{r}R'(r)}{R(r)} = -\lambda \quad (\lambda > 0). \tag{30.128}$$

Relations (30.128) yield two ordinary differential equations

$$T''(t) + \lambda a^2 T(t) = 0, \tag{30.129}$$

$$R''(r) + \frac{1}{r} R'(r) + \lambda R(r) = 0, \tag{30.130}$$

$$R(r_0) = 0, \quad |R(0)| < +\infty. \tag{30.131}$$

The condition $|R(0)| < +\infty$ stems from the natural requirement that the solution $u(r, t)$ be bounded at the membrane's centre, i.e., at $r = 0$. The problem (30.130-131) has the obvious solution $R(r) \equiv 0$, which is no good.

We thus come to the eigenvalue problem of finding those values of λ for which there exist nontrivial solutions of the problem (30.130-131), and the solutions themselves.

We write the equation (30.130) in the form

$$r^2R''(r) + rR'(r) + (\lambda r^2 - 0)R(r) = 0.$$

This is a Bessel equation with $\nu = 0$ (see Chap. 19). Its general solution is

$$R(r) = C_1J_0(\sqrt{\lambda}r) + C_2N_0(\sqrt{\lambda}r).$$

It follows from the condition $|R(0)| < +\infty$ that $C_2 = 0$ (Neumann function $N_0(\sqrt{\lambda}r) \underset{r\to+0}{\longrightarrow} \infty$). Thus, $R(r) = J_0(\sqrt{\lambda}r)$. The boundary condition $R(r_0) = 0$ gives $J_0(\sqrt{\lambda}r_0) = 0$, whence it follows that the number $\sqrt{\lambda}r_0$ must be one of the zeros of the Bessel function $J_0(x)$, i.e., $\sqrt{\lambda}r_0 = \mu_k$, where μ_k is a zero of the function $J_0(x)$. It is well known that the function $J_0(x)$ has an infinite set of positive zeros

$$0 < \mu_1 < \mu_2 < \ldots < \mu_n < \ldots.$$

From this we obtain the eigenvalues

$$\lambda_n = \left(\frac{\mu_n}{r_0}\right)^2 \quad (n = 1, 2, \ldots),$$

and the corresponding eigenfunctions

$$R_n(r) = J_0\left(\frac{\mu_n}{r_0}r\right) \quad (n = 1, 2, \ldots)$$

for the problem (30.130-131).

At $\lambda = \lambda_n$ the general solution of (30.129) has the form

$$T_n(t) = A_n \cos\frac{a\mu_n}{r_0}t + B_n \sin\frac{a\mu_n}{r_0}t.$$

The function

$$u_n(r, t) = \left(A_n \cos \frac{a\mu_n}{r_0} t + B_n \sin \frac{a\mu_n}{r_0} t\right) J_0 \left(\frac{\mu_n}{r_0} r\right)$$

will be a solution to (30.124) satisfying the boundary condition (30.125). It defines standing axisymmetrical waves of the round membrane.

We seek the solution $u(r, t)$ of the original problem (30.124-126) in the form of the formal series

$$u(r, t) = \sum_{n=1}^{\infty} \left(A_n \cos \frac{a\mu_n}{r_0} t + B_n \sin \frac{a\mu_n}{r_0} t\right) J_0 \left(\frac{\mu_n}{r_0} r\right). \qquad (30.132)$$

The coefficients A_n and B_n are found from the initial conditions

$$u(r, 0) = f(r) = \sum_{n=1}^{\infty} A_n J_0 \left(\frac{\mu_0}{r_0} r\right), \qquad (30.133)$$

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = F(r) = \sum_{n=1}^{\infty} \frac{a\mu_n}{r_0} B_n J_0 \left(\frac{\mu_n}{r_0} r\right), \qquad (30.134)$$

i.e., we arrive at the expansion of the given functions $f(x)$ and $F(r)$ into series in Bessel functions.

It can easily be verified that at $m \neq n$ the functions $J_0(\mu_m r/r_0)$ and $J_0(\mu_n r/r_0)$ are orthogonal with weight r on $[0, r_0]$. It is common knowledge that any function $\Phi(r) \in C^2(0, r_0)$ satisfying the boundary conditions of the problem can be expanded into an absolutely and uniformly convergent Fourier-Bessel series

$$\Phi(r) = \sum_{n=1}^{\infty} c_n J_0 \left(\frac{\mu_n}{r_0} r\right),$$

where

$$c_n = \frac{\int_0^{r_0} r\Phi(r) J_0 \left(\frac{\mu_n}{r_0} r\right) dr}{\int_0^{r_0} r J_0^2 \left(\frac{\mu_n}{r_0} r\right) dr} \qquad (n = 1, 2, \ldots).$$

When the initial conditions $f(r)$ and $F(r)$ are sufficiently smooth, we can from this obtain the formulas for the Fourier-Bessel coefficients of

the functions $f(r)$ and $F(r)$

$$A_n = \frac{\int_0^{r_0} r\, f(r)\, J_0\left(\frac{\mu_n}{r_0} r\right) dr}{\int_0^{r_0} r J_0^2\left(\frac{\mu_n}{r_0} r\right) dr} \qquad (n = 1, 2, \ldots).$$

$$B_n = \frac{r_0}{a\mu_n} \frac{\int_0^{r_0} r\, F(r)\, J_0\left(\frac{\mu_n}{r_0} r\right) dr}{\int_0^{r_0} r J_0^2\left(\frac{\mu_n}{r_0} r\right) dr}$$

Substituting the found values of A_n and B_n into (30.132), we obtain the solution of the problem (30.124-126), if the series (30.132) converges uniformly, just like the series that are obtained from it by double termwise differentiation with respect to each argument t and r.

30.11 Application of Laplace Transforms to Solution of Mixed Problems

It is required to find the solution $u(x, t)$ of the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad 0 < x < l, \tag{30.135}$$

satisfying the initial conditions

$$u(x, 0) = \varphi_0(x), \quad \frac{\partial u(x, 0)}{\partial t} = \varphi_1(x), \quad 0 \leqslant x \leqslant l, \tag{30.136}$$

and the boundary conditions

$$u(0, t) = \mu_1(t), \quad u(l, t) = \mu_2(t), \quad t \geqslant 0. \tag{30.137}$$

The independent variable t varies from 0 to $+\infty$. We will therefore apply the Laplace transform in t. Suppose that $u(x, t)$, $\partial u/\partial x$ and $\partial^2 u/\partial x^2$ viewed as functions of t are inverse transforms. Let $U(x, p)$ be the Laplace transform of $u(x, t)$, i.e.,

$$U(x, p) = \int_0^{+\infty} u(x, t) e^{-pt} dt.$$

Assuming the operations of differentiation with respect to x and integration with respect to t in the Laplace transform to be interchangeable, we will get

$$\frac{\partial u}{\partial x} \doteq \frac{\partial}{\partial x}\int_0^{+\infty} u(x, t)e^{-pt}dt = \int_0^{+\infty}\frac{\partial u}{\partial x}e^{-pt}dt = \frac{dU(x, p)}{dx},$$

$$\frac{\partial^2 u}{\partial x^2} \doteq \frac{\partial^2}{\partial x^2}\int_0^{+\infty} u(x, t)e^{-pt}dt = \int_0^{+\infty}\frac{\partial^2 u}{\partial x^2}e^{-pt}dt = \frac{d^2U(x, p)}{dx^2}.$$

Here p is viewed as a parameter and instead of the partial derivatives $\frac{\partial^k}{\partial x^k}$ we write $\frac{d^k}{dx^k}$, $k = 1, 2, \ldots$.

By the rule of differentiation of inverse transforms we have

$$\frac{\partial u}{\partial t} \doteq pU - u(x, 0),$$

$$\frac{\partial^2 u}{\partial t^2} \doteq p^2U - pu(x, 0) - \frac{\partial u(x, 0)}{\partial t}.$$

Taking into account the initial conditions (30.136), we obtain from here

$$\frac{\partial u}{\partial t} \doteq pU - \varphi_0(x), \quad \frac{\partial^2 u}{\partial t^2} \doteq p^2U - p\varphi_0(x) - \varphi_1(x).$$

Let the functions $\mu_1(t)$ and $\mu_2(t)$ in the boundary conditions (30.137) be inverse transforms and let $\mu_1(t) \doteq M_1(p)$, $\mu_2(t) \doteq M_2(p)$. The boundary conditions (30.137) then give

$$U(0, p) = \int_0^{+\infty} u(0, t)e^{-pt}dt = M_1(p),$$

$$U(l, p) = \int_0^{+\infty} u(l, t)e^{-pt}dt = M_2(p).$$

Passing to transforms, we also pass from the problem (30.135-137) for the partial differential equation to the boundary problem for an ordinary differential equation of finding the solution $U(x, p)$ of the equation

$$\frac{d^2U}{dx^2} - \frac{p^2}{a^2}U = -\frac{p}{a^2}\varphi_0(x) - \frac{1}{a^2}\varphi_1(x) \tag{30.138}$$

with the boundary conditions

$$U(0, p) = M_1(p), \quad U(l, p) = M_2(p). \tag{30.139}$$

Let $U(x, p)$ be a solution of the problem (30.138-139). Then the function $u(x, t)$ (the inverse transform for $U(x, p)$) will be a solution of the original problem (30.135-137).

Example. A string of length l is fixed at the ends $x = 0$ and $x = l$. The initial dislocation of the string is given by the formula $u(x, 0) = A \sin \frac{\pi x}{l}$, $A = \text{const}$. There are no initial velocities. Find the dislocation $u(x, t)$ of the string for $t > 0$.

◄ The problem reduces to finding the solution $u(x, t)$ of the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad 0 < x < l \tag{30.140}$$

with the initial conditions

$$u(x, 0) = A \sin \frac{\pi x}{l}, \quad \frac{\partial u(x, 0)}{\partial t} = 0, \quad 0 \leqslant x \leqslant l \tag{30.141}$$

and the boundary conditions

$$u(0, t) = 0, \quad u(l, t) = 0, \quad t \geqslant 0. \tag{30.142}$$

We will apply the Laplace transform in t. Let $U(x, p)$ be the transform of $u(x, t)$

$$U(x, p) = \int_0^{\infty} u(x, t) e^{-pt} dt.$$

Then

$$\frac{\partial u}{\partial x} = \frac{dU}{dx}, \quad \frac{\partial^2 u}{\partial x^2} = \frac{d^2 U}{dx^2}.$$

By the rule of differentiation of inverse transforms with the initial conditions (30.141) taken into account we will have

$$\frac{\partial u}{\partial t} = pU - A \sin \frac{\pi x}{l}, \quad \frac{\partial^2 u}{\partial t^2} = p^2 U - pA \sin \frac{\pi x}{l}.$$

The boundary conditions (30.142) in turn will give

$$U(0, p) = \int_0^{+\infty} u(0, t) e^{-pt} dt = 0,$$

$$U(l, p) = \int_0^{+\infty} u(l, t) e^{-pt} dt = 0.$$

In the space of transforms we thus come to the boundary-value problem for the ordinary differential equation

$$\frac{d^2U}{dx^2} - \frac{p^2}{a^2}U = -\frac{Ap}{a^2}\sin\frac{\pi x}{l}, \tag{30.143}$$

$$U(0, p) = U(l, p) = 0. \tag{30.144}$$

Solving equation (30.143) as a linear equation with constant coefficients, we will find

$$U(x, p) = C_1 e^{\frac{px}{a}} + C_2 e^{-\frac{px}{a}} + \frac{Ap}{p^2 + a^2\pi^2/l^2}\sin\frac{\pi x}{l}.$$

From conditions (30.144) we will obtain $C_1 = C_2 = 0$, so that

$$U(x, p) = \frac{Ap}{p^2 + a^2\pi^2/l^2}\sin\frac{\pi x}{l}.$$

The inverse transform for $U(x, p)$ is the function

$$u(x, t) = A\cos\frac{\pi a t}{l}\sin\frac{\pi x}{l},$$

which will be the solution to problem (30.140-142). ►

In the space of inverse transforms a partial differential equation + initial conditions + boundary conditions

$\downarrow \mathscr{L}$

In the space of transforms an ordinary differential equation + boundary conditions

$\downarrow$

Solution in the space of transforms

$\downarrow \mathscr{L}^{-1}$

Solution of the original problem

Using a similar procedure we solve mixed problems for more general equations of hyperbolic (and parabolic) types.

The course of the solution of a partial differential equation using the Laplace tranform can be represented by the scheme on page 630 (the number of independent variables is $n = 2$).

Exercises

Find the solution of the following original problems:

1. $\dfrac{\partial^2 u}{\partial t^2} = a^2 \dfrac{\partial^2 u}{\partial x^2}$, $t > 0$, $-\infty < x < +\infty$, $u|_{t=0} = e^{-x^2}$, $\left.\dfrac{\partial u}{\partial t}\right|_{t=0} = \cos x$.

2. $\dfrac{\partial^2 u}{\partial t^2} = a^2 \dfrac{\partial^2 u}{\partial x^2}$, $t > 0$, $-\infty < x < +\infty$, $u|_{t=0} = 0$, $\left.\dfrac{\partial u}{\partial t}\right|_{t=0} = \dfrac{1}{1 + x^2}$.

3. A homogeneous infinite string is excited by an initial perturbation in the shape of a semicircle

$$u(x, 0) = \sqrt{1 - x^2}, \quad -1 \leqslant x \leqslant 1.$$

No initial speeds. Draw the position of the string for the times $t = 1/2$, $t = 1$, $t = 2$, taking for simplicity $a = 1$.

Using the Fourier method of separation of variables, solve the following mixed problems:

4. $\dfrac{\partial^2 u}{\partial t^2} = a^2 \dfrac{\partial^2 u}{\partial x^2}$, $t > 0$, $0 < x < l$, $u|_{x=0} = u|_{x=l} = 0$, $u|_{t=0} = \sin \dfrac{3\pi}{l} x$, $\left.\dfrac{\partial u}{\partial t}\right|_{t=0} = 0$. 5. $\dfrac{\partial^2 u}{\partial t^2} = a^2 \dfrac{\partial^2 u}{\partial x^2}$, $t > 0$, $0 < x < l$, $u|_{x=0} = u|_{x=l} = 0$, $u|_{t=0} = 0$, $\left.\dfrac{\partial u}{\partial t}\right|_{t=0} = \sin \dfrac{\pi}{l} x$. 6. $\dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial^2 u}{\partial x^2}$, $t > 0$, $0 < x < \pi$, $u|_{x=0} = u|_{x=\pi} = 0$, $u|_{t=0} = \sin x$, $\left.\dfrac{\partial u}{\partial t}\right|_{t=0} = \sin x$. 7. $\dfrac{\partial^2 u}{\partial t^2} = a^2 \dfrac{\partial^2 u}{\partial x^2}$, $t > 0$, $0 < x < l$, $u|_{x=0} = u|_{x=l} = 0$, $u|_{t=0} = 0$, $\left.\dfrac{\partial u}{\partial t}\right|_{t=0} =$

$$\begin{cases} x, & 0 \leqslant x < \dfrac{l}{2}, \\ l - x, & \dfrac{l}{2} \leqslant x \leqslant l. \end{cases}$$

8. $\dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial^2 u}{\partial x^2} + \sin \pi x$, $t > 0$, $0 < x < 1$, $u|_{x=0} = u|_{x=1} = 0$, $u|_{t=0} = 0$, $\left.\dfrac{\partial u}{\partial t}\right|_{t=0} = 0$. 9. $\dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial^2 u}{\partial x^2} + (4t - 8) \sin 2x$,

$t > 0,\ 0 < x < \pi,\ u|_{x=0} = u|_{x=\pi} = 0,\ u|_{t=0} = \left.\frac{\partial u}{\partial t}\right|_{t=0} = 0.$

Solve the mixed problems:

10. $\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$, $t > 0,\ 0 < x < l,\ u|_{x=0} = 0,\ u|_{x=l} = t,\ u|_{t=0} = 0,$ $\left.\frac{\partial u}{\partial t}\right|_{t=0} = 0.$ 11. $\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$, $t > 0,\ 0 < x < l,\ u|_{x=0} = 0,$ $\left.\frac{\partial u}{\partial x}\right|_{x=l} = A = \text{const},\ u|_{t=0} = 0,\ \left.\frac{\partial u}{\partial t}\right|_{t=0} = 0.$

Answers

1. $u(x, t) = \frac{1}{2}[e^{-(x-at)^2} + e^{-(x+at)^2}] + \frac{\sin at}{a}\cos x.$

2. $u(x, t) = \frac{1}{2a}[\tan^{-1}(x + at) - \tan^{-1}(x - at)].$ 4. $u(x, t) = \sin\frac{3\pi}{l}x\cos\frac{3\pi a}{l}t.$

5. $u(x, t) = \frac{l}{\pi a}\sin\frac{\pi}{l}x\sin\frac{\pi a}{l}t.$ 6. $u(x, t) = (\sin t + \cos t)\sin x.$

7. $u(x, t) = \frac{4l^2}{a\pi^3}\sum_{k=1}^{n}\frac{(-1)^{k+1}}{(2k-1)^3}\sin\frac{(2k-1)\pi}{l}x\cdot\sin\frac{(2k-1)\pi a}{l}t.$

8. $u(x, t) = \frac{1}{\pi^2}(1 - \cos\pi t)\sin\pi x.$ 9. $u(x, t) = \left(2\cos 2t - \frac{\sin 2t}{2} + t - 2\right)\sin 2x.$

10. $u(x, t) = \frac{xt}{l} + \frac{2l}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\sin\frac{\pi n}{l}x\sin\frac{\pi n}{l}t.$

11. $u(x, t) = Ax + \frac{8Al}{\pi^2}\sum_{k=0}^{\infty}\frac{(-1)^{k+1}}{(2k+1)^2}\sin\frac{(2k+1)\pi}{2l}x\cos\frac{(2k+1)\pi}{2l}t.$

Hint: Put $u = v + Ax$, where v is a new unknown function.

Chapter 31

Parabolic Equations

31.1 Heat Equation

Parabolic partial differential equations of the second order occur in dealing with processes of heat conduction and diffusion.

We will now derive the equation that governs the distribution of temperature in a heat conducting material. We will denote by $u(x, y, z, t)$ the temperature in a medium at a point $M(x, y, z)$ at a time t. Considering the medium isotropic, we denote by $\varrho(M)$ its density, by $c(M)$ its specific thermal capacity, and by $k(M)$ the thermal conductivity at M. Inside the body heat may be produced or absorbed (e.g., due to a chemical reaction). We will denote by $F(M, t)$ the density of heat sources at the point M at a time t.

Next we calculate the heat balance in an arbitrary volume V during a time interval $(t, t + dt)$. Let S be the boundary of V and $\mathbf{n}$ be the external normal to S. If the temperature of the body is nonuniformly distributed, then heat fluxes arise in it. According to the Fourier law through the surface S into the volume V comes the following amount of heat:

$$Q_1 = \iint_S k\frac{\partial u}{\partial n} ds\, dt = \iint_S (k \,\mathbf{grad}\, u, \mathbf{n}^0) ds\, dt,$$

where $\mathbf{n}^0$ is the unit vector of the external normal to S.

To the integral on the right we apply the Ostrogradsky-Gauss theorem. We will have

$$Q_1 = \iiint_V \operatorname{div}(k \,\mathbf{grad}\, u) dv\, dt.$$

The input of the heat sources inside V is

$$Q_2 = \iiint_V F(x, y, z, t) dv\, dt.$$

Suppose that during the time interval $(t, t + dt)$ the temperature in V increased by

$$\Delta u = u(M, t + dt) - u(M, t) \approx \frac{\partial u}{\partial t} dt.$$

The physics of the process dictates that for this change to occur it is necessary to have the heat input

$$Q_3 = \iiint\limits_V c\varrho \frac{\partial u}{\partial t}\, dv\, dt.$$

From energy conservation law we have

$$Q_3 = Q_1 + Q_2,$$

therefore

$$\iiint\limits_V \left[\operatorname{div}(k \operatorname{\mathbf{grad}} u) + F - c\varrho \frac{\partial u}{\partial t}\right] dv\, dt = 0.$$

Volume V being arbitrary, we obtain the equation

$$c\varrho \frac{\partial u}{\partial t} = \operatorname{div}(k \operatorname{\mathbf{grad}} u) + F(M, t). \tag{31.1}$$

If the medium is homogeneous, i.e., if c, ϱ, and k are constants, then equation (31.1) becomes

$$\frac{\partial u}{\partial t} = a^2 \Delta u + f, \tag{31.2}$$

where

$$a^2 = \frac{k}{c\varrho}, \qquad f = \frac{F}{c\varrho}, \qquad \Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

Equation (31.2) is called the *heat equation*. Similarly, we derive the equation of diffusion of particles.

As in the case of the equation of oscillations, for heat conduction process to be described completely, it is necessary to specify an initial distribution of the temperature $u(M, t)$ in a medium (initial condition) and the conditions on the boundary of the medium (boundary conditions).

We will confine ourselves to the heat equation with one spatial variable

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t)$$

(heat propagation in linear bodies).

31.2 Cauchy Problem for Heat Equation

We consider the homogeneous heat equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2},$$

which corresponds to $f(x, t) \equiv 0$, i.e., to the case of no sources. We formulate the Cauchy problem as follows: find a function $u(x, t)$ satisfying the equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} \qquad (t > 0, \qquad -\infty < x < +\infty) \tag{31.3}$$

and the initial condition

$$u\big|_{t=0} = \varphi(x) \qquad (-\infty < x < +\infty). \tag{31.4}$$

Physically, the problem is to find the temperature of a homogeneous infinite rod at any moment of time $t > 0$, when its temperature $\varphi(t)$ at $t = 0$ is known. The sides of the rod are taken to be thermally insulated, so that no heat leaves the rod through it.

We suppose that:

(1) $u(x, t)$ and $\varphi(x)$ are sufficiently smooth and as $x^2 + t^2 \to +\infty$ they decrease so fast that there exist Fourier transforms

$$v(\xi, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u(x, t)\, e^{-i\xi x} dx, \tag{31.5}$$

$$\tilde{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \varphi(x)\, e^{-i\xi x} dx, \tag{31.6}$$

(2) differentiation operations are legitimate:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\partial u}{\partial t} e^{-i\xi x} dx = \frac{dv}{dt},$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{\partial^2 u}{\partial x^2} e^{-i\xi x} dx = -\xi^2 v(\xi, t).$$

If then we apply Fourier transform to both sides of equation (31.3) and to the condition (31.4), we will thus come to the Cauchy problem for the ordinary differential equation

$$\frac{dv}{dt} + \xi^2 a^2 v = 0, \tag{31.7}$$

$$v\big|_{t=0} = \tilde{\varphi}(\xi), \tag{31.8}$$

the quantity ξ here plays the role of a parameter.

The solution to the problem (31.7-8) has the form

$$v(\xi, t) = \tilde{\varphi}(\xi)\, e^{-\xi^2 a^2 t}. \tag{31.9}$$

We have earlier established (see Chap. 27) that

$$\mathcal{F}[e^{-\alpha x^2}] = \frac{1}{\sqrt{2\alpha}} e^{-\xi^2/4\alpha},$$

where $\mathcal{F}[f]$ is the Fourier transform of $f(x)$.

Putting $t = 1/4a^2\alpha$, we obtain

$$e^{-\xi^2 a^2 t} = \mathcal{F}\left[\frac{1}{a\sqrt{2t}} e^{-x/4a^2 t}\right].$$

On the right of (31.9) we have thus the product of the Fourier transforms of the functions $\varphi(x)$ and $\exp(-x^2/4a^2t)/a\sqrt{2t}$.

Using now the convolution theorem according to which

$$\mathcal{F}[f_1 * f_2] = \sqrt{2\pi}\,\mathcal{F}[f_1]\,\mathcal{F}[f_2],$$

we can represent (31.9) as

$$v(\xi, t) = \tilde{\varphi}(\xi) e^{-\xi^2 a^2 t} = \frac{1}{\sqrt{2\pi}} \mathcal{F}\left[\varphi(x) * \frac{1}{a\sqrt{2t}} e^{-x^2/4a^2 t}\right]. \tag{31.10}$$

The left side of this is the Fourier transform (in x) of the required function $u(x, t)$, so that we can rewrite (31.10) as

$$\mathcal{F}[u(x, t)] = \frac{1}{2a\sqrt{\pi t}} \mathcal{F}[\varphi(x) * e^{-x^2/4a^2 t}].$$

Using the expression for the convolution of the functions $\varphi(x)$ and $\exp(-x^2/4a^2t)$, we will obtain

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}} \int_{-\infty}^{+\infty} \varphi(\lambda) e^{-\frac{(x-\lambda)^2}{4a^2 t}} d\lambda, \quad t > 0. \tag{31.11}$$

This formula is the solution of the original problem (31.3-4) and is called the *Poisson intergal.*

Indeed, we can show that for any continuous and bounded function $\varphi(x)$ the function $u(x, t)$ given by (31.11) has derivatives of any order in x and t for $t > 0$ and obeys equation (31.3) for $t > 0$ and all x.

We now show that the function (31.11) for $\varphi(x) \in C(-\infty, +\infty)$ obeys the initial condition $u|_{t=0} = \varphi(x)$, $-\infty < x < +\infty$.

We put

$$\frac{x - \lambda}{2a\sqrt{t}} = \mu,$$

then

$$\lambda = x - 2a\sqrt{t}\,\mu, \qquad d\lambda = -2a\sqrt{t}\,d\mu,$$

so that

$$u(x, t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \varphi(x - 2a\sqrt{t}\mu)\, e^{-\mu^2} d\mu.$$

Hence as $t \to +0$ we find

$$u(x, 0) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \varphi(x)\, e^{-\mu^2} d\mu = \varphi(x),$$

since

$$\int_{-\infty}^{+\infty} e^{-\mu^2} d\mu = \sqrt{\pi}.$$

We will now formulate another result.

Theorem 31.1. *In the class of bounded functions* $u(x, t)$: $\{|u(x, t)| < M, -\infty < x < +\infty, t > 0\}$ *the solution of the Cauchy problem* (31.3-4) *is unique and continuously dependent on the initial function.*

Example. Find the solution of the Cauchy problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad t > 0, \quad -\infty < x < +\infty, \tag{31.3'}$$

$$u|_{t=0} = e^{-x^2/2}, \quad -\infty < x < \infty. \tag{31.4'}$$

◄ Using the Poisson formula (31.11), we find for $\varphi(x) = e^{-x^2/2}$, $a = 1$ that

$$u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{+\infty} e^{-\lambda^2/2}\, e^{-(x-\lambda)^2/4t}\, d\lambda. \tag{31.12}$$

We transform the integral on the right-hand side of (31.12)

$$\int_{-\infty}^{+\infty} e^{-\lambda^2/2}\, e^{-(x-\lambda)^2/4t} d\lambda = \int_{-\infty}^{+\infty} e^{-\lambda^2/2 - x^2/4t + x\lambda/2t - \lambda^2/4t}\, d\lambda$$

$$= e^{-x^2/2(1+2t)} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{1+2t}{2t}\left(\lambda - \frac{x}{1+2t}\right)^2} d\lambda. \tag{31.13}$$

We now change the variable

$$\frac{\sqrt{1+2t}}{\sqrt{2t}}\left(\lambda - \frac{x}{1+2t}\right) = \xi.$$

The integral on the right of (31.13) will then become

$$\frac{\sqrt{2t}}{\sqrt{1+2t}}\int_{-\infty}^{+\infty} e^{-\xi^2/2}\,d\xi = \frac{2\sqrt{\pi t}}{\sqrt{1+2t}}.$$

(We have here used the fact that $\int_{-\infty}^{+\infty} e^{-\xi^2/2}\,d\xi = \sqrt{2\pi}$). Therefore, we will obtain from (31.13)

$$\int_{-\infty}^{+\infty} e^{-\lambda^2/2}\,e^{-(x-\lambda)^2/4t}\,d\lambda = \frac{2\sqrt{\pi t}}{\sqrt{1+2t}}\,e^{-x^2/2(1+2t)}.$$

The solution $u(x, t)$ of the original problem will thus be given by

$$u(x, t) = \frac{1}{\sqrt{1+2t}}\,e^{-\frac{x^2}{2(1+2t)}}, \qquad t > 0. \quad ► \tag{31.14}$$

Remark. It follows from the Poisson formula (31.11) that heat instantaneously propagates along the rod. Indeed, let the initial temperature $\varphi(x)$ be positive for $\alpha \leqslant x \leqslant \beta$ and zero outside the interval. The subsequent distribution of temperature will then be

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}}\int_{\alpha}^{\beta} \varphi(\lambda)\,e^{-\frac{(x-\lambda)^2}{4a^2t}}\,d\lambda, \qquad t > 0.$$

This suggests that for arbitrarily small $t > 0$ and arbitrarily large $|x|$ we have $u(x, t) > 0$. This is explained by inaccuracy of the theoretical premises in deriving the heat equation, namely the neglect of the inertia of molecular motion. Nevertheless, the heat equation provides reasonable agreement with experiment. A more rigorous description of heat transfer processes is given by the so-called transfer equations.

Fundamental solution of heat equation. The function $G(x, t; \lambda) = \exp[-(x-\lambda)^2/4a^2t]/2a\sqrt{\pi t}$ in the Poisson formula (31.11) is called the *fundamental solution* of the heat equation. Viewed as a function of x, t, the function $G(x, t; \lambda)$ obeys the equation $u_t = a^2u_{xx}$, which can be verified by a direct check. The fundamental solution has an important physical meaning, which is associated with the notion of the *thermal pulse.*

Suppose that the initial distribution $\varphi(x)$ of temperature is

$$\varphi(x) = \varphi_\varepsilon(x) = \begin{cases} \dfrac{1}{2\varepsilon} & \text{if} \quad |x - x_0| < \varepsilon, \\ 0 & \text{if} \quad |x - x_0| > \varepsilon. \end{cases}$$

Using (31.11), we then find the temperature distribution $u(x, t)$, $t > 0$, in the rod

$$u(x, t) = \frac{1}{2\varepsilon} \frac{1}{2a\sqrt{\pi t}} \int_{x_0-\varepsilon}^{x_0+\varepsilon} e^{-\frac{(x-\lambda)^2}{4a^2 t}} d\lambda. \tag{31.15}$$

By the mean value theorem of integral calculus, we obtain

$$\int_{x_0-\varepsilon}^{x_0+\varepsilon} e^{-\frac{(x-\lambda)^2}{4a^2 t}} d\lambda = 2\varepsilon e^{-\frac{(x-\bar{\lambda})^2}{4a^2 t}},$$

where $\bar{\lambda} \in [x_0 - \varepsilon, x_0 + \varepsilon]$, so that by (31.15)

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}} e^{-\frac{(x-\bar{\lambda})^2}{4a^2 t}}.$$

Passing to the limit as $\varepsilon \to 0$, we will have

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}} e^{-\frac{(x-x_0)^2}{4a^2 t}} = G(x, t; x_0).$$

This means that the function $G(x, t; x_0)$ represents the temperature distribution in the rod for $t > 0$, if at $t = 0$ and $x = x_0$ there was an infinite peak of temperature (as $\varepsilon \to 0$ the function $\varphi_\varepsilon(x) \to +\infty$), and elsewhere in the rod the temperature was zero. Such an initial distribution of temperatures can approximately be realized as follows: at $t = 0$ we bring up for an instant to a point $x = x_0$ on the rod a narrow flame of exceedingly high temperature (a heat pulse of density $c\varrho$). This initial distribution of temperature is described by the so-called Dirac *δ-function*, denoted as $\delta(x - x_0)$.

Not a function in a conventional sense, δ-function is defined formally by

(1) $\delta(x - x_0) = \begin{cases} 0 & \text{at} \quad x \neq x_0 \\ +\infty & \text{at} \quad x = x_0 \end{cases},$

(2) $\int_\alpha^\beta (x - x_0)dx = 1$ in any interval (α, β) that contains the point x_0.

The main property of δ-function is that for any continuous function $f(x)$

$$\int_\alpha^\beta f(x)\delta(x - x_0)dx = \begin{cases} f(x_0) & \text{if} \quad x_0 \in (\alpha, \beta) \\ 0 & \text{if} \quad x_0 \overline{\in} (\alpha, \beta) \end{cases}.$$

The fundamental solution $G(x, t; x_0)$ is thus a solution of the heat equation for the infinite rod with the initial distribution of temperature

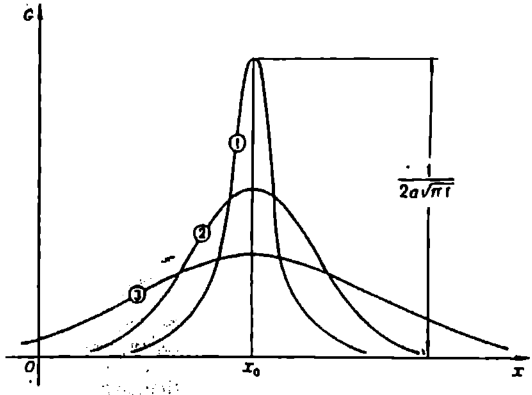


Fig. 31.1

$\varphi(x) = \delta(x - x_0)$. The graph of $G(x, t; x_0)$ for various values of $t > 0$ has the form shown in Fig. 31.1.

Curves *1*, *2*, *3* correspond to the moments of time $0 < t_1 < t_2 < t_3$, respectively. The figure shows how the temperature in the rod levels off after the heat pulse. The solution

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}} \int_{-\infty}^{+\infty} \varphi(\lambda) e^{-\frac{(x-\lambda)^2}{4a^2 t}} d\lambda$$

of the heat conduction problem in the infinite rod with the initial condition $u|_{t=0} = \varphi(x)$ can be treated as a result of the superposition of temperatures at the point x at time t due to heat pulses of intensity $\varphi(\lambda)$ at the point λ applied at $t = 0$ and distributed uniformly over the rod.

31.3 Heat Propagation in a Finite Rod

If a rod is of finite length l and occupies the segment $0 \leqslant x \leqslant l$ on the x-axis, then to formulate the problem of heat propagation in such a rod, in addition to the equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t) \tag{31.16}$$

and the initial condition

$$u|_{t=0} = \varphi(x) \tag{31.17}$$

it is necessary to know its temperature conditions at the ends of the rod $x = 0$ and $x = l$, i.e., to specify its boundary conditions. The boundary conditions may differ depending on the temperature conditions at the rod's ends. We consider three main types of the boundary conditions.

(1) At the ends of the rod we have the temperature

$$u(0, t) = \mu_1(t); \qquad u(l, t) = \mu_2(t),$$

where $\mu_1(t)$, $\mu_2(t)$ are functions defined for the time span $0 \leqslant t \leqslant T$ during which we study the process.

(2) At the ends of the rod we have the values of the derivative

$$\left.\frac{\partial u}{\partial x}\right|_{x=0} = \nu_1(t), \qquad \left.\frac{\partial u}{\partial x}\right|_{x=l} = \nu_2(t).$$

These conditions occur if we know the value of the heat flux Q through the end face of the rod. For example, if at $x = l$ the quantity $Q(l, t)$ is given, then

$$Q(l, t) = -k\left.\frac{\partial u}{\partial x}\right|_{x=l},$$

hence $\partial u/\partial x|_{x=l} = \nu_2(t)$, where $\nu_2(t) = -Q(l, t)/k$. If $\nu_1(t)$ or $\nu_2(t)$ are identically zero, then we say that the corresponding end is thermally insulated.

(3) At the ends of the rod we have the following linear relations between the function and its derivative:

$$\left.\frac{\partial u}{\partial x}\right|_{x=0} = \lambda[u(0, t) - \theta(t)],$$

$$\left.\frac{\partial u}{\partial x}\right|_{x=l} = -\lambda[u(l, t) - \theta(t)],$$

where $\theta(t)$ is a known function—the ambient temperature, λ is the heat exchange coefficient. This boundary condition corresponds to heat exchange according to Newton's law on the boundary of a body with an environment whose temperature is $\theta(t)$. Using two expressions for the heat flux through the cross section $x = l$: $Q = h(u - \theta)$ and $Q = -k\dfrac{\partial u}{\partial x}$, we will arrive at the following form of the third boundary condition:

$$\left.\frac{\partial u}{\partial x}\right|_{x=l} = -\lambda[u(l, t) - \theta(t)], \qquad \lambda = \frac{h}{k}.$$

For the cross section $x = 0$ of the rod the third boundary condition has the form $\left.\dfrac{\partial u}{\partial x}\right|_{x=0} = +\lambda[u(0, t) - \theta(t)]$, since for the heat flux $-k\dfrac{\partial u}{\partial n}$ at $x = 0$ we have $\dfrac{\partial u}{\partial n} = -\dfrac{\partial u}{\partial x}$ (the external normal to the rod at the end $x = 0$ has the opposite direction to the x-axis).

The above basic problems by no means exhaust the possible boundary problems for the equation $u_t = a^2 u_{xx} + f(x, t)$. For example, at the different ends of the rod conditions of different types can be specified.

We will confine ourselves to the first mixed problem for the heat equation.

The problem is formulted as follows: find the solution $u(x, t)$ of the equation (31.16) in the domain $0 < x < l$, $t > 0$, $u(x, t) \in C^2\ \{0 < x < l, t > 0\}$, meeting the initial condition (31.17) for $0 \leqslant x \leqslant l$, and the boundary conditions

$$u|_{x=0} = \mu_1(t), \qquad u|_{x=l} = \mu_2(t), \qquad t \geqslant 0. \tag{31.18}$$

We consider that the function $u(x, t)$ is continuous in the closed domain $\overline{D}\ \{0 \leqslant x \leqslant l,\ 0 \leqslant t \leqslant T\}$ (Fig. 31.2), which requires that the functions $\varphi(x)$, $\mu_1(t)$, $\mu_2(t)$ were continuous and that the conditions $\varphi(0) = \mu_1(0)$, $\varphi(l) = \mu_2(l)$ be met.

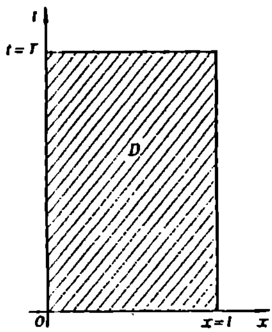


Fig. 31.2

Remark. Just as for hyperbolic equations, the function $u(x, t)$ is sought only for $0 < x < l$ and $t > 0$, but not at $t = 0$ and not at $x = 0$ and $x = l$, where the values of the function $u(x, t)$ are predetermined by the initial and boundary conditions.

We formulate the maximum principle.

Theorem 31.2. *If the function $u(x, t) \in C(\overline{D})$ obeys the heat equation $\dfrac{\partial u}{\partial t} = a^2 \dfrac{\partial^2 u}{\partial x^2}$ in the domain $D\ \{0 < x < l,\ 0 < t \leqslant T\}$, then the maximum and minimum values of $u(x, t)$ are achieved either at $t = 0$ or on the boundaries $x = 0$ or $x = l$.*

The physical meaning of the theorem is obvious: if the temperature of a body on the boundary or at $t = 0$ does not exceed some value M, then inside the body (no sources!) there may be no temperature higher than M. The following theorems are consequences of the maximum principle.

Theorem 31.3 (on uniqueness). *The solution of the problem* (31.16-18) *in the rectangle* D $\{0 < x < l,\ 0 < t \leqslant T\}$ *is unique.*

Theorem 31.4. *The solution of the problem* (31.16-18) *is continuously dependent on the initial and boundary conditions.*

31.4 Fourier Method for Heat Equation

We now turn to solution of the first mixed problem involving the heat equation. We find the solution $u(x, t)$ of the equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t), \qquad t > 0, \qquad 0 < x < l, \tag{31.19}$$

meeting the initial condition

$$u|_{t=0} = \varphi(x), \qquad 0 \leqslant x \leqslant l \tag{31.20}$$

and the boundary conditions

$$u|_{x=0} = \mu_1(t), \qquad u|_{x=l} = \mu_2(t), \qquad t \geqslant 0. \tag{31.21}$$

(1) We will begin with the simplest problem of finding the solution $u(x, t)$ of the homogeneous equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{31.22}$$

satisfying the initial condition

$$u|_{t=0} = \varphi(x), \tag{31.23}$$

and the zero (homogeneous) boundary conditions

$$u|_{x=0} = 0, \qquad u|_{x=l} = 0. \tag{31.24}$$

We will seek the nontrivial solutions of equation (31.22) satisfying the boundary conditions (31.24) in the form

$$u(x, t) = T(t)X(x). \tag{31.25}$$

Substituting $u(x, t)$ in the form (31.25) into (31.22), we will get

$$T'(t)X(x) = a^2T(t)X''(x)$$

or

$$\frac{T'(t)}{a^2T(t)} = \frac{X''(x)}{X(x)} = -\lambda, \tag{31.26}$$

whence we obtain two ordinary differential equations

$$T'(t) + a^2\lambda T(t) = 0, \tag{31.27}$$

$$X''(x) + \lambda X(x) = 0. \tag{31.28}$$

To obtain the nontrivial solution $u(x, t)$ of the form (31.25) satisfying the boundary conditions (31.24) it is necessary to find the nontrivial solutions of (31.28) meeting the boundary conditions

$$X(0) = 0, \qquad X(l) = 0. \tag{31.29}$$

Thus to find $X(x)$ we will have to solve an eigenvalue problem of finding the values of λ at which there exist nontrivial solutions of the problem

$$X''(x) + \lambda X(x) = 0, \tag{31.30}$$

$$X(0) = 0, \qquad X(l) = 0. \tag{31.31}$$

This problem has been considered in Chap. 30. It has been shown that only at

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2 \quad (n = 1, 2, \ldots,)$$

there exist nontrivial solutions $X_n(x) = \sin(n\pi x/l)$ of the problem (31.30-31).

At $\lambda = \lambda_n$ the general solution of (31.27) has the form

$$T_n(t) = a_n e^{-\left(\frac{n\pi a}{l}\right)^2 t}$$

where a_n are arbitrary constants. The functions

$$u_n(x, t) = a_n e^{-\left(\frac{n\pi a}{l}\right)^2 t} \sin\frac{n\pi}{l} x \qquad (n = 1, 2, \ldots)$$

obey equation (31.22) and the boundary conditions (31.24).

We set up the formal series

$$u_n(x, t) = \sum_{n=1}^{\infty} a_n e^{-\left(\frac{n\pi a}{l}\right)^2 t} \sin\frac{n\pi}{l} x. \tag{31.32}$$

We require that the function $u(x, t)$ given by (31.32) satisfied the initial condition $u|_{t=0} = \varphi(x)$ to obtain

$$\varphi(x) = \sum_{n=1}^{\infty} a_n \sin\frac{n\pi}{l} x. \tag{31.33}$$

Series (31.33) represents the expansion of the given function $\varphi(x)$ in a sine Fourier series in the interval $(0, l)$. The coefficients a_n of the expansion are given by the well-known formulas (see Chap. 16).

$$a_n = \frac{2}{l}\int_0^l \varphi(x) \sin\frac{n\pi}{l} x\,dx \qquad (n = 1, 2, \ldots). \tag{31.34}$$

Suppose that $\varphi(x) \in C^2[0, 1]$ and $\varphi(0) = \varphi(l) = 0$. Series (31.33) with coefficients given by (31.34) will then converge to $\varphi(x)$ absolutely and uniformly. Since for $t \geqslant 0$

$$0 < e^{-\left(\frac{n\pi a}{l}\right)^2 t} \leqslant 1,$$

then series (31.32) for $t \geqslant 0$ also converges absolutely and uniformly. Therefore, the function $u(x, t)$ is the sum of series (31.32), which is continuous in the domain $0 < x < l$, $t > 0$ and satisfies the initial and boundary conditions. It remains to show that the function $u(x, t)$ satisfies the equation (31.22) in the domain $0 < x < l$, $t > 0$. It is sufficient to show that the series obtained from (31.32) by termwise differentiation with respect to t once and termwise differentiation with respect to x twice also converges absolutely and uniformly for $0 < x < l$, $t > 0$. But this follows from the fact that for any $t > 0$ we have

$$0 < \frac{n^2\pi^2a^2}{l^2} e^{-\left(\frac{n\pi a}{l}\right)^2 t} < 1,$$

$$0 < \frac{n^2\pi^2}{l^2} e^{-\left(\frac{n\pi a}{l}\right)^2 t} < 1$$

if n is sufficiently large.

The uniqueness of the solution of problem (31.22-24) and the continuous dependence of the solution on the initial function $\varphi(x)$ have already been established above. The problem is thus stated correctly for $t > 0$; conversely, for $t < 0$ the problem is not correct.

Remark. The equation $\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}$ is asymmetric in t; if we substitute $-t$ for t, we will then obtain an equation of another form, i.e., $\frac{\partial u}{\partial t} = -a^2 \frac{\partial^2 u}{\partial x^2}$, whereas the wave equation $\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$ is symmetric in time.

The equation $\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}$ describes irreversible processes. So, we can predict the form of u in a period of time t, but we cannot with confidence tell what u was like a while before. This difference between prediction and prehistory is typical for the parabolic equation and is not the case, for the wave equation, which enables us to look with ease both into the future and into the past.

Example. Find the distribution of temperature in a uniform rod of length π, if the initial temperature of the rod is $u|_{t=0} = \sin x$ $(0 \leqslant x \leqslant \pi)$ and at the ends of the rod the temperature is zero.

◄ We have the equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad t > 0, \qquad 0 < x < \pi \tag{31.35}$$

with the initial condition

$$u|_{t=0} = \sin x, \qquad 0 \leqslant x \leqslant \pi \tag{31.36}$$

and the boundary conditions

$$u|_{x=0} = 0, \qquad u|_{x=\pi} = 0 \qquad t \geqslant 0. \tag{31.37}$$

Using the Fourier method, we seek the nontrivial solutions of equation (31.35) meeting the boundary conditions (31.37) in the form

$$u(x, t) = T(t)X(x). \tag{31.38}$$

Substituting $u(x, t)$ in the form (31.38) into (31.35) and separating the variables, we will obtain

$$\frac{T'(t)}{a^2 T(t)} = \frac{X''(x)}{X(x)} = -\lambda,$$

hence

$$T'(t) + a^2\lambda\, T(t) = 0 \tag{31.39}$$

and

$$X''(x) + \lambda X(x) = 0, \tag{31.40}$$

$$X(0) = X(\pi) = 0. \tag{31.41}$$

The eigenvalues of the problem (31.40-41) are $\lambda_n = n^2 (n = 1, 2, \ldots)$, the eigenfunctions are $X_n(x) = \sin nx$. When $\lambda = \lambda_n$, the general solution (31.39) is $T_n(t) = a_n e^{-a^2 n^2 t}$, so that

$$u_n(x, t) = a_n e^{-a^2 n^2 t} \sin nx.$$

We will seek the solution of the problem (31.35-37) as the series

$$u(x, t) = \sum_{n=1}^{\infty} a_n e^{-a^2 n^2 t} \sin nx. \tag{31.42}$$

If we require that the initial condition (31.36) be met, we will get

$$u(x, 0) = \sin x = \sum_{n=1}^{\infty} a_n \sin nx,$$

hence $a_1 = 1$, $a_k = 0$ $(k = 2, 3, \ldots)$. Therefore, the solution of the original

problem will be the function

$$u(x,\ t) = e^{-a^2 t}\sin x. \ ►$$

(2) Now consider the problem of finding the solution $u(x,\ t)$ of the inhomogeneous equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x,\ t), \qquad t > 0, \qquad 0 < x < l, \tag{31.43}$$

meeting the initial condition

$$u|_{t=0} = \varphi(x), \qquad 0 \leqslant x \leqslant l, \tag{31.44}$$

and the homogeneous boundary conditions

$$u|_{x=0} = 0, \qquad u|_{x=l} = 0, \qquad t \geqslant 0. \tag{31.45}$$

We now suppose that $f(x,\ t)$ is continuous, has the continuous derivative $\frac{\partial f}{\partial x}$, and for all $t > 0$ we have $f(0,\ t) = f(l,\ t) = 0$.

We will seek the solution of the problem (31.43-45) in the form

$$u(x,\ t) = v(x,\ t) + w(x,\ t), \tag{31.46}$$

where $v(x,\ t)$ is the solution of the problem

$$\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 v}{\partial x^2} + f(x,\ t), \tag{31.47}$$

$$v|_{t=0} = 0, \tag{31.48}$$

$$v|_{x=0} = 0, \qquad v|_{x=l} = 0 \tag{31.49}$$

and the function $w(x,\ t)$ is the solution of the problem

$$\frac{\partial w}{\partial t} = a^2 \frac{\partial^2 w}{\partial x^2}, \tag{31.50}$$

$$w|_{t=0} = \varphi(x), \tag{31.51}$$

$$w|_{x=0} = w|_{x=l} = 0, \tag{31.52}$$

The problem (31.50-52) has been considered in (1). We will seek the solution $v(x,\ t)$ of the problem (31.47-49) as the series

$$v(x,\ t) = \sum_{n=1}^{\infty} T_n(t) \sin \frac{n\pi}{l} x \tag{31.53}$$

in eigenfunctions $\{\sin n\pi x/l\}$ of the boundary problem $X''(x) + \lambda X(x) = 0$, $X(0) = X(l) = 0$. Substituting $v(x,\ t)$ in the form (31.53) into (31.47), we get

$$\sum_{n=1}^{\infty} \left(T_n'(t) + \frac{n^2\pi^2 a^2}{l^2} T_n(t) \right) \sin \frac{n\pi}{l} x = f(x,\ t). \tag{31.54}$$

We expand the function $f(x, t)$ into a sine Fourier series

$$f(x, t) = \sum_{n=1}^{\infty} f_n(t) \sin \frac{n\pi}{l} x, \tag{31.55}$$

where

$$f_n(t) = \frac{2}{l} \int_0^l f(\xi, t) \sin \frac{n\pi}{l} \xi \, d\xi. \tag{31.56}$$

Comparing the expansions (31.54) and (31.55) of $f(x, t)$ into a Fourier series, we obtain

$$T_n'(t) + \left(\frac{n\pi a}{l}\right)^2 T_n(t) = f_n(t) \qquad (n = 1, 2, \ldots). \tag{31.57}$$

Using the initial condition for $v(x, t)$:

$$v(x, 0) = 0 = \sum_{n=1}^{\infty} T_n(0) \sin \frac{n\pi}{l} x, \qquad 0 \leqslant x \leqslant l,$$

we find that

$$T_n(0) = 0 \qquad (n = 1, 2, \ldots). \tag{31.58}$$

The solution of (31.57) for the initial conditions (31.58) has the form

$$T_n(t) = \int_0^t f_n(\tau) e^{-\left(\frac{n\pi a}{l}\right)^2 (t-\tau)} d\tau \qquad (n = 1, 2, \ldots).$$

Substituting these expressions for $T_n(t)$ into (31.53), we will find the solution $v(x, t)$ of the problem (31.47-49)

$$v(x, t) = \sum_{n=1}^{\infty} \left[\int_0^t f_n(\tau) e^{-\left(\frac{n\pi a}{l}\right)^2 (t-\tau)} d\tau \right] \sin \frac{n\pi}{l} x. \tag{31.59}$$

The function $u(x, t) = w(x, t) + v(x, t)$ will be a solution of the original problem (31.43-45).

(3) Consider the problem of finding in the domain $\{0 < x < l,\ t > 0\}$ the solution $u(x, t)$ of the equation

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t) \tag{31.60}$$

with the initial condition

$$u|_{t=0} = \varphi(x) \tag{31.61}$$

and the inhomogeneous boundary conditions

$$u|_{x=0} = \mu_1(t), \qquad u|_{x=l} = \mu_2(t) \tag{31.62}$$

We cannot directly apply the Fourier method since the conditions (31.62) are inhomogeneous.

We now introduce a new unknown function $v(x, t)$ by setting

$$u(x, t) = v(x, t) + \omega(x, t),$$

where

$$\omega(x, t) = \mu_1(t) + [\mu_2(t) - \mu_1(t)]\frac{x}{l}.$$

The solution of the problem (31.60-62) will then come down to the solution of the problem (31.43-45) considered in (2) for the function $v(x, t)$.

Exercises

1. Given an infinite uniform rod, show that if the initial temperature is

$$\varphi(x) = u_0 e^{-\sigma^2 x^2}, \qquad -\infty < x < +\infty \quad (u_0 > 0,\ \sigma > 0 = \text{const}),$$

then at any moment $t > 0$ the temperature of the rod will be

$$u(x, t) = \frac{u_0}{\sqrt{1 + 4a^2\sigma^2 t}}\, e^{-\frac{\sigma^2 x^2}{1+4a^2\sigma^2 t}}.$$

2. The ends of a rod of length π are maintained at zero temperature. The initial temperature is given by $u(x, 0) = 2\sin 3x$. Find the temperature of the rod for any moment of time $t > 0$.
3. The ends of a rod of length l are maintained at zero temperature. The initial temperature of the rod is given by $u(x, 0) = 3\sin \pi x/l - 5\sin 2\pi x/l$. Find the temperature of the rod for any moment of time $t > 0$.
4. The ends of a rod of length l are maintained at zero temperature. The initial distribution of temperature is

$$\varphi(x) = \begin{cases} \dfrac{2u_0}{l}x, & 0 \leqslant x \leqslant \dfrac{l}{2}, \\[2mm] \dfrac{2u_0}{l}(l - x), & \dfrac{l}{2} < x \leqslant l \end{cases} \quad (u_0 = \text{const}).$$

Find the temperature of the rod for any moment of time $t > 0$.

Answers

2. $u(x, t) = 2e^{-9a^2 t}\sin 3x$. 3. $u(x, t) = 3e^{-\frac{a^2\pi^2}{l^2}t}\sin\frac{\pi}{l}x - 5e^{-\frac{4a^2\pi^2}{l^2}t}\sin\frac{2\pi}{l}x$.

4. $u(x, t) = \dfrac{u_0}{2} - \dfrac{4u_0}{\pi^2}\displaystyle\sum_{k=0}^{\infty}\frac{\cos\frac{(4k+2)\pi}{l}x}{(2k+1)^2}\, e^{-\frac{(4k+2)^2\pi^2 a^2}{l^2}t}$.

Chapter 32

Elliptic Equations

32.1 Definition. Formulation of Boundary Problems

Elliptic equations turn up in studies of stationary, i.e., time-independent, processes of various physical nature. The simplest equation of elliptical type is the Laplace equation

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0. \qquad (32.1)$$

This equation characterizes gravitational and electrostatic potential in points of free space and the temperature of a homogeneous isotropic medium with an established heat motion.

In the case of a function $u = u(x, y)$ of two independent variables x, y, the Laplace equation has the form

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \qquad (32.2)$$

It lies at the heart of the theory of analytic functions of a complex variable. Its solutions are the real and imaginary parts of the function $f(z) = u(x, y) + iv(x, y)$ that are analytic in some domain D. For a function of one argument $u = u(x)$ we have

$$\Delta u \equiv \frac{d^2 u}{dx^2} = 0. \qquad (32.3)$$

Solutions of (32.3) are the functions $u = C_1 x + C_2$, where C_1 and C_2 are arbitrary constants.

Definition. The function $u(x, y, z)$ is *harmonic* in the domain $\Omega \subset R^3$, if $u \in C^2(\Omega)$ and obeys in Ω the Laplace equation (32.1).

Let a domain Ω be bounded by a surface Σ (Fig. 32.1). A typical problem for the Laplace equation is to find a function $u(M)$, $M \in \Omega$, that is harmonic in Ω and satisfying in Σ the boundary condition that can be taken in any of the following forms:

(1) $u|_{\Sigma} = f_1(p)$, $p \in \Sigma$—the first boundary problem, or Dirichlet problem,

(2) $\partial u/\partial n|_\Sigma = f_2(p)$, $p \in \Sigma$—the second boundary problem, or the Neumann problem,

(3) $(\partial u/\partial n + hu)|_\Sigma = f_3(p)$, $p \in \Sigma$—the third boundary problem. Here f_1, f_2, f_3, h are predetermined functions, $\partial u/\partial n$ is the derivative along the external normal to the surface Σ.

The geometrical meaning of the Dirichlet problem for the unidimensional Laplace equation is trivial. Unidimensional harmonic functions $u = C_1 x + C_2$ are direct lines and the Dirichlet problem reduces to the following: to draw a straight line through two points $A(x_1, u_1)$ and $B(x_2, u_2)$ (Fig. 32.2).

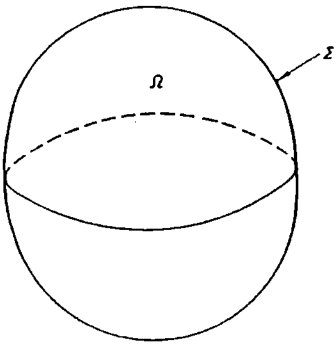


Fig. 32.1

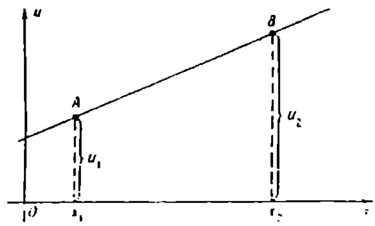


Fig. 32.2

Depending on where we seek the solution of the problem (inside the region bounded by a surface Σ or outside of Σ) we distinguish internal and external boundary problems for the equation $\Delta u = 0$.

Another representative of elliptic equations is the *Poisson equation*

$$\Delta u = g(x, y, z),$$

which corresponds to an equilibrium state under the action of an external force with a density proportional to $g(x, y, z)$.

Let us take another example, namely, the wave equation

$$\Delta u - \frac{1}{a^2} u_{tt} = 0. \tag{32.4}$$

We will seek solution of (32.4) in the form

$$u(x, y, z, t) = v(x, y, z)e^{i\omega t}. \tag{32.5}$$

Substituting (32.5) into (32.4) gives

$$e^{i\omega t}\Delta v + \frac{\omega^2}{a^2} v e^{i\omega t} = 0,$$

hence

$$\Delta v + k^2 v = 0,$$

where $k^2 = \omega^2/a^2$.

We have thus obtained for $v(x, y, z)$ the elliptic equation

$$\Delta v + k^2 v = 0,$$

which is called the *Helmholtz equation.*

As with the Laplace equation, typical boundary problems for the Poisson and Helmholtz equations are those of the first, second and third types.

32.2 Fundamental Solutions of Laplace Equation

The commonest coordinates are the Cartesian, cylindrical and spherical coordinates. The Laplace operator in the Cartesian coordinates x, y, z is given by

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2},$$

in the cylindrical coordinates r, φ, z, by

$$\Delta u \equiv \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\partial^2 u}{\partial z^2},$$

in the spherical coordinates r, θ, φ, by

$$\Delta u \equiv \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial u}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial \varphi^2}.$$

Of great interest are solutions of the Laplace equation featuring a spherical or cylindrical symmetry, i.e., ones that only depend on one variable r.

In spherical coordinates, we find that the solution $u = u(r)$ that has a spherical symmetry is found from the ordinary differential equation

$$\frac{d}{dr}\left(r^2\frac{du}{dr}\right) = 0.$$

Integrating this equation gives

$$u = \frac{C_1}{r} + C_2 \qquad (C_1, C_2 = \text{const}).$$

Putting, say, $C_1 = 1$ and $C_2 = 0$, we will get the function

$$u_0(r) = \frac{1}{r},$$

which is called the *fundamental solution of the Laplace equation in space.*

The function $u_0 = 1/r$ obeys the equation $\Delta u = 0$ everywhere except for the point $r = 0$, where u_0 becomes infinite. If we look at the field of a point charge e placed at the origin, then the potential of the field will be $u = e/r$.

Using the cylindrical coordinates, we find that the solution $u = u(r)$, which shows a cylindrical or circular symmetry (for two independent variables), is found from the ordinary differential equation

$$\frac{d}{dr}\left(r\frac{du}{dr}\right) = 0.$$

We integrate this to obtain

$$u = C_1 \ln r + C_2.$$

Choosing $C_1 = -1$ and $C_2 = 0$, we will have

$$u_0(r) = \ln\frac{1}{r}.$$

The function $u_0(r)$ is called the *fundamental solution of the Laplace equation in a plane*. This function satisfies the Laplace equation everywhere, except for the point $r = 0$, where $u = \ln(1/r)$ becomes infinite.

32.3 Green's Formulas

We will proceed from the Ostrogradsky-Gauss formula (see Chap. 24)

$$\oint\!\!\!\oint_{\Sigma} (\mathbf{a}, \mathbf{n}^0)\, d\sigma = \iiint_{\Omega} \operatorname{div} \mathbf{a}\, d\Omega. \qquad (32.6)$$

We put $\mathbf{a} = v \operatorname{\mathbf{grad}} u$ and assume that u and v have continuous second derivatives in Ω and are continuous together with their first derivatives in $\overline{\Omega} = \Omega \cup \Sigma$, i.e.,

$$u,\ v \in C^2(\Omega) \cap C^1(\overline{\Omega}).$$

We have

$$(\mathbf{a}, \mathbf{n}^0) = v(\operatorname{\mathbf{grad}} u, \mathbf{n}^0) = v\frac{\partial u}{\partial n},$$

$$\operatorname{div}\mathbf{a} = (\nabla, v\nabla u) = (\nabla v, \nabla u) + v\nabla^2 u = (\operatorname{\mathbf{grad}} u, \operatorname{\mathbf{grad}} v) + v\Delta u.$$

Therefore, from (32.6)

$$\iiint_{\Omega} [(\operatorname{\mathbf{grad}} u, \operatorname{\mathbf{grad}} v) + v\Delta u]\, d\Omega = \iint_{\Sigma} v\frac{\partial u}{\partial n}\, d\sigma. \qquad (32.7)$$

This is the *first Green's formula.*

Interchanging u and v in (32.7) gives

$$\iiint_{\Omega} [(\text{grad } v, \text{grad } u) + u\,\Delta v]\,d\Omega = \iint_{\Sigma} u\,\frac{\partial v}{\partial n}\,d\sigma. \tag{32.8}$$

Substituting (32.7) from (32.8) termwise, we find

$$\iiint_{\Omega} [(u\,\Delta v - v\,\Delta u]\,d\Omega = \iint_{\Sigma} \left(u\,\frac{\partial v}{\partial n} - v\,\frac{\partial u}{\partial n}\right) d\sigma. \tag{32.9}$$

This is the *second Green's formula.*

Lastly, setting in (32.7) $v = u$, we will obtain

$$\iiint_{\Omega} [(\text{grad } u)^2 + u\,\Delta u]\,d\Omega = \iint_{\Sigma} u\,\frac{\partial u}{\partial n}\,d\sigma. \tag{32.10}$$

This is the *third Green's formula.*

Everywhere here $\mathbf{n}$ is the external normal to the surface Σ, where Σ is a smooth or piecewise smooth closed surface.

The boundary Σ of the domain Ω can consist of several closed surfaces. In this case, the surface integrals on the right-hand sides of Green's formulas must be taken over all the surfaces that bound the domain Ω.

32.4 Basic Integral Green's Formula

We have established that the function $v(M) = 1/r_{MM_0}$, where $r_{MM_0} = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$ is the distance between the points $M(x, y, z)$ and $M_0(x_0, y_0, z_0)$ satisfies the Laplace equation $\Delta u = 0$ for all M, $M \neq M_0$. Let Ω be a domain in R^3 with a boundary Σ, and $u(M) \in C^2(\Omega) \cap C^1(\overline{\Omega})$. Consider the function $v = 1/r_{MM_0}$, where M_0 is some internal point in Ω. Since this function is continuous at point $M_0 \in \Omega$, we may not directly apply the second Green's formula

$$\iiint_{\Omega} (u\,\Delta v - v\,\Delta u)\,d\Omega = \iint_{\Sigma} \left(u\,\frac{\partial v}{\partial n} - v\,\frac{\partial u}{\partial n}\right) d\sigma$$

to u and v. Consider the domain $\Omega \setminus K_\varepsilon$ with the boundary $\Sigma \cup \Sigma_\varepsilon$, which we obtain if from the domain Ω we exclude a sphere K_ε of radius ε and centre at M_0 and surface Σ_ε (Fig. 32.3). Applying the second of Green's formulas to the functions $u(M)$ and $v(M) = 1/r_{MM_0}$ in the domain $\Omega \setminus K_\varepsilon$, we obtain

$$\iiint_{\Omega \setminus K_\varepsilon} \left(u\Delta\left(\frac{1}{r}\right) - \frac{1}{r}\,\Delta u\right) d\Omega = \iint_{\Sigma} \left(u\,\frac{\partial}{\partial n}\left(\frac{1}{r}\right) - \frac{1}{r}\,\frac{\partial u}{\partial n}\right) d\sigma$$
$$+ \iint_{\Sigma_\varepsilon} u\,\frac{\partial}{\partial n}\left(\frac{1}{r}\right) d\sigma - \iint_{\Sigma_\varepsilon} \frac{1}{r}\,\frac{\partial u}{\partial n}\,d\sigma. \tag{32.11}$$

We transform the second integral on the right of (32.11). Calculating the derivative along the external normal to the region $\Omega\backslash K_\varepsilon$ on the surface Σ_ε (Fig. 32.4), we find that

$$\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\bigg|_{\Sigma_\varepsilon} = -\frac{\partial}{\partial r}\left(\frac{1}{r}\right)\bigg|_{\Sigma_\varepsilon} = \frac{1}{\varepsilon^2}.$$

Using the mean value theorem for the integral over the surface Σ_ε we find

$$\iint\limits_{\Sigma_\varepsilon} u\frac{\partial}{\partial n}\left(\frac{1}{r}\right)d\sigma = \frac{1}{\varepsilon^2}\iint\limits_{\Sigma_\varepsilon} u\,d\sigma = \frac{1}{\varepsilon^2}4\pi\varepsilon^2 u^* = 4\pi u^*, \qquad (32.12)$$

where u^* is the mean value of $u(M)$ over the surface Σ_ε.

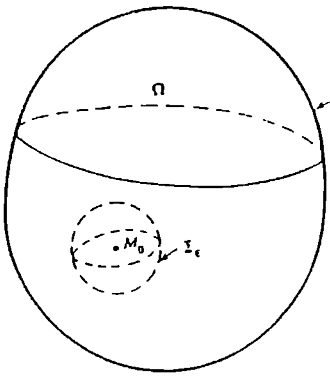


Fig. 32.3

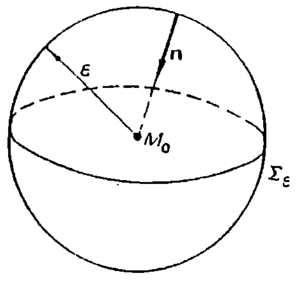


Fig. 32.4

We now transform the third integral on the right of (32.11)

$$\iint\limits_{\Sigma_\varepsilon}\frac{1}{r}\frac{\partial u}{\partial n}d\sigma = \frac{1}{\varepsilon}\iint\limits_{\Sigma_\varepsilon}\frac{\partial u}{\partial n}d\sigma = \frac{1}{\varepsilon}4\pi\varepsilon^2\left(\frac{\partial u}{\partial n}\right)^* = 4\pi\varepsilon\left(\frac{\partial u}{\partial n}\right)^*, \qquad (32.13)$$

where $(\partial u/\partial n)^*$ is the mean value of the normal derivative $\partial u/\partial n$ over the sphere Σ_ε. Substituting (32.12) and (32.13) into (32.11) and considering that $\Delta(1/r_{MM_0}) \equiv 0$ in $\Omega\backslash K_\varepsilon$, we will have

$$\iiint\limits_{\Omega\backslash K_\varepsilon}\left(-\frac{1}{r}\right)\Delta u\,d\Omega = \iint\limits_{\Sigma}\left(u\frac{\partial}{\partial n}\left(\frac{1}{r}\right) - \frac{1}{r}\frac{\partial u}{\partial n}\right)d\sigma$$
$$+ 4\pi u^* - 4\pi\varepsilon\left(\frac{\partial u}{\partial n}\right)^*. \qquad (32.14)$$

Let us now the radius ε tend to zero. We then obtain:

(1) $\lim_{\varepsilon\to 0} u^* = u(M_0)$, since $u(M)$ is a continuous function and u^* is its mean value over the sphere of radius ε with centre at point M_0;

(2) $\lim_{\varepsilon\to 0} 4\pi\varepsilon(\partial u/\partial n)^* = 0$, since the normal derivative

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\cos\beta + \frac{\partial u}{\partial z}\cos\gamma$$

is bounded in the neighbourhood of the point M_0, because the first derivatives of $u(M)$ are continuous inside Ω;

(3) by the definition of the improper integral we have

$$\lim_{\varepsilon\to 0}\iiint\limits_{\Omega\setminus K_\varepsilon}\left(-\frac{1}{r}\right)\Delta u\, d\Omega = -\iiint\limits_{\Omega}\frac{\Delta u}{r}\, d\Omega.$$

The limiting process in (32.14) when $\varepsilon \to 0$ leads us to the *basic integral Green's formula*

$$u\,(M_0) = \frac{1}{4\pi}\iint\limits_{\Sigma}\left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\right)d\sigma - \frac{1}{4\pi}\iiint\limits_{\Omega}\frac{\Delta u}{r}d\Omega. \tag{32.15}$$

We can thus represent any function $u(M) \in C^2(\Omega)\cap C^1(\overline{\Omega})$ as the sum of three integrals

$$-\frac{1}{4\pi}\iiint\limits_{(\Omega)}\frac{\Delta u}{r}\,d\Omega,\quad \frac{1}{4\pi}\iint\limits_{\Sigma}\frac{1}{r}\frac{\partial u}{\partial n}\,d\sigma,\quad -\frac{1}{4\pi}\iint\limits_{\Sigma}u\frac{\partial}{\partial n}\left(\frac{1}{r}\right)d\sigma,$$

which are referred to as the volume potential, the potential of the simple layer and the potential of the double layer, respectively.

If $u(M)$ is a harmonic function in Ω, then $\Delta u \equiv 0$ and the formula (32.15) assumes the form

$$u(M_0) = \frac{1}{4\pi}\iint\limits_{\Sigma}\left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\right)d\sigma \tag{32.16}$$

(the point M_0 lies inside Ω).

This is the *principal formula of the theory of harmonic functions*. It tells us that the value of a harmonic function at any internal point in a domain Ω is expressed through the value of the function and its normal derivative on the boundary Σ of the domain Ω. It follows from (32.16) that any harmonic function $u(M)$ in Ω is the sum of the potentials of a simple and double layers.

For the Laplace equation in the plane the fundamental solution has the form $v_0 = \ln(1/r)$. Using exactly the same arguments, we obtain the

basic integral formula for a harmonic function of two arguments

$$u(M_0) = \frac{1}{2\pi} \oint_{\Gamma} \left[\left(\ln \frac{1}{r} \right) \frac{\partial u}{\partial n} - u \frac{\partial}{\partial n} \left(\ln \frac{1}{r} \right) \right] ds. \quad (32.17)$$

Here Γ is the boundary of a domain D, n is the normal to the boundary (Fig. 32.5). Any harmonic function $u(x, y)$ in D is thus the sum of two potentials

$$\frac{1}{2\pi} \oint_{\Gamma} \left(\ln \frac{1}{r} \right) \frac{\partial u}{\partial n} ds, \qquad -\frac{1}{2\pi} \oint_{\Gamma} u \frac{\partial}{\partial n} \left(\ln \frac{1}{r} \right) ds.$$

The first of this is the logarithmic potential of the simple layer and the second one is the logarithmic potential of the double layer.

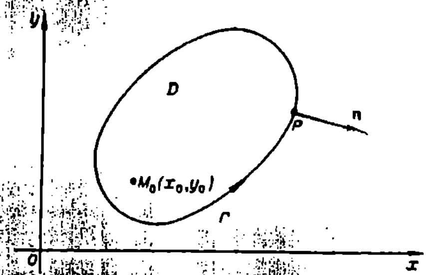


Fig. 32.5

32.5 Properties of Harmonic Functions

Theorem 32.1. *If a function $u(M)$ is harmonic in a domain Ω and continuous, together with its first derivatives, in $\bar{\Omega} = \Omega \cup \Sigma$, then its normal derivative $\partial u/\partial n$ on the boundary Σ of Ω meets the condition*

$$\iint_{\Sigma} \frac{\partial u}{\partial n} d\sigma = 0. \quad (32.18)$$

◄ Applying Green's second formula

$$\iiint_{\Omega} (u \Delta v - v \Delta u) d\Omega = \iint_{\Sigma} \left(u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) d\sigma$$

to a harmonic function $u(M)$ and to the function $v(M) \equiv 1$, we will obtain

$$0 = \iint_{\Sigma} \frac{\partial u}{\partial n} d\sigma. \ ►$$

This property suggests that there are no sources inside Ω.

Theorem 32.2. *If there exists a solution to the Neumann problem for the Laplace equation, then it is defined up to a constant addend.*

◄ To prove this statement we additionally suppose that $u \in C^1(\overline{\Omega})$. Let there be two solutions $u_1(M)$ and $u_2(M)$. Then

$$\Delta u_1 \equiv 0, \qquad \Delta u_2 \equiv 0$$

and

$$\left.\frac{\partial u_1}{\partial n}\right|_\Sigma = f(P), \qquad \left.\frac{\partial u_2}{\partial n}\right|_\Sigma = f(P).$$

The difference $u = u_1 - u_2$ will be a solution of the problem

$$\Delta u = 0, \qquad \left.\frac{\partial u}{\partial n}\right|_\Sigma = 0.$$

Using Green's third formula

$$\iiint_\Omega ((\text{grad}\, u)^2 + u\,\Delta u)\,d\Omega = \iint_\Sigma u\,\frac{\partial u}{\partial n}\,d\sigma$$

we will have for such a function $u(M)$

$$\iiint_\Omega (\text{grad}\, u)^2 d\Omega = 0 \qquad \text{or}$$

$$\iiint_\Omega \left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2\right) d\Omega = 0.$$

The integrand being continuous and nonnegative, we get

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 \equiv 0$$

or

$$\frac{\partial u}{\partial x} \equiv 0, \qquad \frac{\partial u}{\partial y} \equiv 0, \qquad \frac{\partial u}{\partial z} \equiv 0 \qquad \text{in} \qquad \Omega,$$

so that $u(x, y, z) \equiv 0$, and hence $u_1(M) - u_2(M) \equiv \text{const.}$ ►

We stress that in the Neumann problem the function $f(P)$, $P \in \Sigma$, must meet the condition

$$\iint_\Sigma f(P)\,d\sigma = 0. \tag{32.19}$$

If this condition is not satisfied, the Neumann problem has no solutions.

Theorem 32.3 (on the mean value of a harmonic function). *If a function $u(M)$ is harmonic inside a ball of radius R with centre at a point M_0 and continuous together with the first derivatives up to the boundary $\Sigma_R^{M_0}$, then the value of $u(M)$ at the centre M_0 of the sphere $\Sigma_R^{M_0}$ is equal to*

the arithmetic mean of all the values $u(M)$ on that sphere, i.e.,

$$u(M_0) = \frac{1}{4\pi R^2} \iint\limits_{\Sigma_R^{M_0}} u(P)\,d\sigma, \tag{32.20}$$

where R is the radius of the sphere.

◀ We will apply the basic integral formula of the theory of harmonic functions

$$u(M_0) = \frac{1}{4\pi} \iint\limits_{\Sigma} \left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\right) d\sigma$$

to the sphere $\Sigma_R^{M_0}$. For this sphere (Fig. 32.6) $r = M_0P = R$, $\partial/\partial n = \partial/\partial r$, therefore

$$u(M_0) = \frac{1}{4\pi} \iint\limits_{\Sigma_R^{M_0}} \left(\frac{1}{R}\frac{\partial u}{\partial n} + u\frac{1}{R^2}\right) d\sigma. \tag{32.21}$$

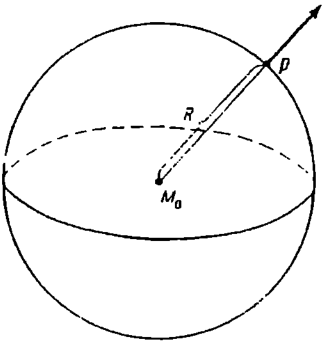


Fig. 32.6

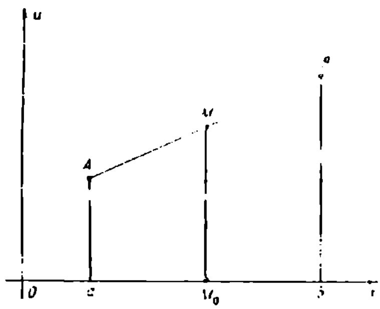


Fig. 32.7

By Theorem 32.1 we have $\iint\limits_{\Sigma_R^{M_0}} \frac{\partial u}{\partial n}\,d\sigma = 0$ and formula (32.21) yields

$$u(M_0) = \frac{1}{4\pi R^2} \iint\limits_{\Sigma_R^{M_0}} u(P)\,d\sigma. \quad ▶$$

For the unidimensional Laplace equation $d^2u/dx^2 = 0$ this theorem is a theorem on the median of a trapezoid: the length of the segment M_0M is equal to half the sum of the lengths of the segments aA and bB (Fig. 32.7).

Theorem 32.4 (on the extrema of harmonic functions). *Let a function $u(M)$ be harmonic in a domain Ω and not identically equal to a constant. Then it has no local extrema inside Ω.*

◄ We will give the proof by contradiction. Suppose that $u(M)$ has at point $M_0 \in \Omega$ a local maximum, i.e.,

$$u(M_0) > u(M) \tag{32.22}$$

at all points M of the ball of a sufficiently small radius with centre at M_0. By the theorem on the mean value of a harmonic function

$$u(M_0) = \frac{1}{4\pi R^2} \iint_{\Sigma_R^{M_0}} u(M)\,d\sigma.$$

($\Sigma_R^{M_0}$ is a sphere within the above ball). By the theorem on the mean value for the double integral we have

$$u(M_0) = \frac{1}{4\pi R^2} 4\pi R^2 u(M_m) = u(M_m).$$

This is at variance with (32.22), and thus proves the theorem. ►

A function $u(M)$, harmonic in a bounded domain Ω and continuous in a closed domain $\overline{\Omega} = \Omega \cup \Sigma$, attains its absolute maximum or minimum on the boundary Σ of Ω (the maximum principle).

From this we can readily deduce the following theorems:

Theorem 32.5 (uniqueness theorem). *The solution of the internal Dirichlet problem*

$$\Delta u = 0$$
$$u|_\Sigma = f(P), \qquad P \in \Sigma,$$

continuous in a closed domain $\overline{\Omega} = \Omega \cup \Sigma$ is unique.

◄ Suppose that we have two solutions of the problem: $u_1(M)$ and $u_2(M)$. Then the difference $u(M) = u_1(M) - u_2(M)$ is harmonic in Ω, continuous in $\overline{\Omega}$ and equal to zero on Σ. By Theorem 32.4, the absolute maximum and minimum values of $u(M)$ in Ω are zero. Consequently, $u(M) = u_1(M) - u_2(M) \equiv 0$ in Ω, i.e., $u_1(M) \equiv u_2(M)$. ►

Theorem 32.6 (on continuous dependence of solution of the first internal boundary problem on boundary values). *Let $u_1(M)$ and $u_2(M)$ be solutions of the problems*

$$\Delta u = 0, \qquad u|_\Sigma = \varphi_1(P),$$

and

$$\Delta u = 0 \qquad u|_\Sigma = \varphi_2(P)$$

that are continuous in $\overline{\Omega} = \Omega \cup \Sigma$. Then, if everywhere on the boundary Σ holds the inequality

$$|\varphi_1(P) - \varphi_2(P)| < \varepsilon, \qquad P \in \Sigma,$$

then throughout Ω *will hold*

$$|u_1(M) - u_2(M)| < \varepsilon, \qquad M \in \Omega.$$

◄ The function $u(M) = u_1(M) - u_2(M)$ is harmonic in Ω and continuous in $\overline{\Omega}$, and $u|_\Sigma = \varphi_1(P) - \varphi_2(P)$. Since $-\varepsilon < \varphi_1(P) - \varphi_2(P) < \varepsilon$, by Theorem 32.4 the absolute maximum and minimum values of $u(M)$ are contained between $-\varepsilon$ and ε. Hence $|u(M)| < \varepsilon$, i.e., $|u_1(M) - u_2(M)| < \varepsilon \; \forall M \in \Omega$. ►

32.6 Solution of the Dirichlet Problem for a Circle Using the Fourier Method

The problem is formulated as follows: find a function $u(r, \varphi)$ that inside a circle K_{r_0} of radius r_0 with centre at the origin obeys the Laplace equation

$$\Delta u = 0, \tag{32.23}$$

that is continuous in the closed region $\overline{K_{r_0}}$ and assumes specified values on the boundary of the circle, i.e.,

$$u|_{r=r_0} = f(\varphi). \tag{32.24}$$

The function $f(\varphi)$ is assumed to be sufficiently smooth and periodic with period 2π.

The desired solution being single-valued, it must be periodic in φ with period 2π, i.e., $u(r, \varphi) = u(r, \varphi + 2\pi)$. And since the solution is continuous in $\overline{K_{r_0}}$, it is bounded in K_{r_0}.

In polar coordinates equation (32.23) has the form

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} = 0. \tag{32.25}$$

We will seek particular solutions of equation (32.25) in the form

$$u(r, \varphi) = R(r)\Phi(\varphi). \tag{32.26}$$

Substituting $u(r, \varphi)$ in the form (32.26) into (32.25) multiplied by r^2, we will get

$$\Phi(\varphi) r \frac{d}{dr}\left(r\frac{dR}{dr}\right) + R(r)\Phi''(\varphi) = 0$$

or

$$\frac{r\dfrac{d}{dr}\left(r\dfrac{dR}{dr}\right)}{R(r)} = -\frac{\Phi''(\varphi)}{\Phi(\varphi)} = \lambda,$$

hence

$$\Phi''(\varphi) + \lambda\Phi(\varphi) = 0, \tag{32.27}$$

$$r^2R''(r) + rR'(r) - \lambda R(r) = 0. \tag{32.28}$$

From the condition $u(r, \varphi) = u(r, \varphi + 2\pi)$ it follows that $\Phi(\varphi) = \Phi(\varphi + 2\pi)$. We then find from (32.27) that $\lambda = n^2$ $(n = 0, 1, 2, \ldots)$ so that

$$\Phi_n(\varphi) = A_n \cos n\varphi + B_n \sin n\varphi \qquad (n = 0, 1, 2, \ldots).$$

Specifically, $\Phi_0(\varphi) = A_0 = \text{const}$.

We look for the solution $R(r)$ of equation (32.28) (Euler equation) in the form $R(r) = r^\sigma$.

Assuming $R(r) = r^\sigma$ we find from (32.28) (at $\lambda = n^2$)

$$\sigma(\sigma - 1) + \sigma - n^2 = 0.$$

Thus, $\sigma^2 - n^2 = 0$, $\sigma = \pm n$ $(n > 0)$, and hence

$$R_n(r) = a_n r^n + b_n r^{-n} \qquad (n > 0).$$

At $n = 0$ we find from (32.28)

$$R_0(r) = a_0 + b_0 \ln r.$$

To solve the internal Dirichlet problem we have to put $b_n = 0$, $n = 0, 1, 2, \ldots$ (since $r^{-n} \to \infty$ and $\ln r \to \infty$ as $r \to 0 + 0$), i.e., we have to take $R_n(r) = a_n r^n$ $(n = 1, 2, \ldots)$, $R_0(r) = a_0$.

We will now look for the solution of the internal Dirichlet problem in the form

$$u(r, \varphi) = A_0 + \sum_{n=1}^{\infty} r^n(A_n \cos n\varphi + B_n \sin n\varphi), \tag{32.29}$$

where the coefficients A_n, B_n are found from the boundary condition (32.24).

At $r = r_0$ we have

$$u(r_0, \varphi) = f(\varphi) = A_0 + \sum_{n=1}^{\infty} r_0^n (A_n \cos n\varphi + B_n \sin n\varphi). \tag{32.30}$$

We write the expansion of $f(\varphi)$ into a Fourier series

$$f(\varphi) = \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} (\alpha_n \cos n\varphi + \beta_n \sin n\varphi), \tag{32.31}$$

where

$$\alpha_0 = \frac{1}{\pi} \int_0^{2\pi} f(t)dt, \qquad \alpha_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt\, dt,$$

$$\beta_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt\, dt \qquad (n = 1, 2, \ldots). \tag{32.32}$$

A comparison of (32.30) and (32.31) yields

$$A_0 = \frac{\alpha_0}{2}, \qquad A_n = \frac{\alpha_n}{r_0^n}, \qquad B_n = \frac{\beta_n}{r_0^n} \qquad (n = 1, 2, \ldots).$$

The formal solution of the internal Dirichlet problem can thus be represented as a sum $u(r, \varphi)$ of the series

$$u(r, \varphi) = \frac{\alpha_0}{2} + \sum_{n=1}^{\infty} \left(\frac{r}{r_0}\right)^n (\alpha_n \cos n\varphi + \beta_n \sin n\varphi), \tag{32.33}$$

where the coefficients $\alpha_0, \alpha_1, \beta_1, \ldots, \alpha_n, \beta_n, \ldots$ are given by (32.32).

For $r < r_0$ the series (32.33) can be differentiated with respect to r and φ any number of times, and hence the function $u(r, \varphi)$ from (32.33) satisfies the equation $\Delta u = 0$.

If we assume that $f(\varphi)$ is continuous and differentiable, then series (32.33) will converge uniformly for $r \leqslant r_0$, and so $u(r, \varphi)$ will be continuous on the boundary of the circle and will meet all the conditions of the problem.

The solution of the external Dirichlet problem should be sought for as a series

$$u(r, \varphi) = \sum_{n=0}^{\infty} \frac{1}{r^n} (A_n \cos n\varphi + B_n \sin n\varphi), \tag{32.34}$$

where the coefficients A_n, B_n are to be found from the boundary condition $u|_{r=r_0} = f(\varphi)$.

For the annulus $r_1 < r < r_2$ formed by two concentric circles with centre at point O and radii r_1 and r_2 (Fig. 32.8) we will seek the solution of the

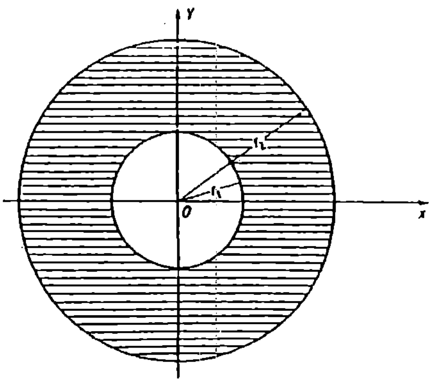


Fig. 32.8

problem in the form of a series

$$u(r, \varphi) = \sum_{n=1}^{\infty} \left(A_n r^n + \frac{C_n}{r^n}\right) \cos n\varphi$$

$$+ \sum_{n=1}^{\infty} \left(B_n r^n + \frac{D_n}{r^n}\right) \sin n\varphi + A_0 \ln r + B_0, \qquad (32.35)$$

whose coefficients $A_0, B_0, A_n, C_n, B_n, D_n$ $(n = 1, 2, \ldots)$ are to be found from the boundary conditions $u(r_1, \varphi) = f_1(\varphi)$ and $u(r_2, \varphi) = f_2(\varphi)$.

Example. Find a function that is harmonic inside a circle of radius r_0 with centre at the origin, such that $u|_{r=r_0} = 3 + 5\cos\varphi$.

◄ The problem comes down to solving the internal Dirichlet problem for the equation $\Delta u = 0$ with the boundary condition

$$u|_{r=r_0} = 3 + 5\cos\varphi. \qquad (32.36)$$

We will seek for the solution in the form of a series

$$u(r, \varphi) = A_0 + \sum_{n=1}^{\infty} r^n (A_n \cos n\varphi + B_n \sin n\varphi).$$

We have from (32.36)

$$u(r_0, \varphi) = 3 + 5\cos\varphi = A_0 + \sum_{n=0}^{\infty} r_0^n (A_n \cos n\varphi + B_n \sin n\varphi).$$

Since the system of functions $1, \sin\varphi, \cos\varphi, \ldots$ is orthogonal on $[0, 2\pi]$ we obtain

$$A_0 = 3, \qquad r_0 A_1 = 5, \qquad A_n = 0 \quad \forall n \geqslant 2, \qquad B_n = 0 \quad \forall n.$$

The desired solution will be

$$u(r, \varphi) = 3 + \frac{5}{r_0} r \cos\varphi \quad \text{or} \quad u(x, y) = 3 + \frac{5}{r_0} x. \text{ ►}$$

32.7 Poisson Integral

We simplify formula (32.33). Substituting expressions (32.32) for the Fourier coefficients into (32.33) and changing the order of summation and integration, we will get

$$u(r, \varphi) = \frac{1}{\pi} \int_0^{2\pi} f(t) \left\{\frac{1}{2} + \sum_{n=1}^{\infty} \left(\frac{r}{r_0}\right)^2 (\cos nt \cos n\varphi + \sin nt \sin n\varphi\right\} dt$$

$$=\frac{1}{\pi}\int_0^{2\pi} f(t)\left\{\frac{1}{2}+\sum_{n=1}^{\infty}\left(\frac{r}{r_0}\right)^2 \cos n(\varphi-t)\right\} dt. \tag{32.37}$$

We will put for short $r/r_0 = \tau$ and make the following rearrangements ($\tau < 1$):

$$\frac{1}{2}+\sum_{n=1}^{\infty}\tau^n \cos n(\varphi-t)=\frac{1}{2}+\frac{1}{2}\sum_{n=1}^{\infty}\tau^n[e^{in(\varphi-t)}+e^{-in(\varphi-t)}]$$

$$+\frac{1}{2}\left[1+\sum_{n=1}^{\infty}(\tau e^{i(\varphi-t)})^n+\sum_{n=1}^{\infty}(\tau e^{-i(\varphi-t)})^n\right]$$

$$=\frac{1}{2}\left[1+\frac{\tau e^{i(\varphi-t)}}{1-\tau e^{i(\varphi-t)}}+\frac{\tau e^{-i(\varphi-t)}}{1-\tau e^{-i(\varphi-t)}}\right]$$

$$=\frac{1}{2}\,\frac{1-\tau^2}{1-2\tau\cos(\varphi-t)+\tau^2}.$$

Substitution into (32.37) gives

$$u(r,\varphi)=\frac{1}{2\pi}\int_0^{2\pi} f(t)\frac{1-\dfrac{r^2}{r_0^2}}{1-2\dfrac{r}{r_0}\cos(\varphi-t)+\dfrac{r^2}{r_0^2}}\,dt$$

or

$$u(r,\varphi)=\frac{1}{2\pi}\int_0^{2\pi} f(t)\frac{r_0^2-r^2}{r_0^2-2rr_0\cos(\varphi-t)+r^2}\,dt. \tag{32.38}$$

The resulting formula, which is the solution of the first boundary problem for the equation $\Delta u = 0$ inside the circle K_{r_0}, is called the *Poisson integral* and the expression

$$\frac{r_0^2-r^2}{r_0^2-2rr_0\cos(\varphi-t)+r^2}$$

is called the *Poisson kernel*.

It can be shown that if $f(\varphi)$ is only continuous on the circumference K_{r_0}, then the function

$$u(r,\varphi)=\begin{cases}\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} f(t)\frac{r_0^2-r^2}{r_0^2-2rr_0\cos(\varphi-t)+r^2}\,dt & \text{for } r<r_0,\\[2ex] f(\varphi) & \text{for } r=r_0\end{cases}$$

satisfies the equation $\Delta u = 0$ for $r < r_0$ and is continuous on the closed circle $\overline{K}_{r_0}$.

Remark. The solution of the external boundary problem has the form

$$u(r, \varphi) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \frac{r^2 - r_0^2}{r^2 - 2rr_0 \cos(\varphi - t) + r_0^2} dt \qquad (r > r_0).$$

Exercises

Find a function that is harmonic inside a circle of radius r_0 with centre at the origin, such that

1. $u|_{r=r_0} = 2 + 3\sin\varphi$. 2. $u|_{r=r_0} = \sin^2\varphi$. 3. $\left.\frac{\partial u}{\partial r}\right|_{r=r_0} = A\cos\varphi$ (A = const). 4. $\left.\frac{\partial u}{\partial r}\right|_{r=r_0} = 2\sin^2\varphi$.

5. Find a function that is harmonic in the annulus $1 < r < 2$, such that

$$u|_{r=1} = 1 - \cos\varphi, \qquad u|_{r=2} = \sin 2\varphi.$$

6. Find the stationary distribution of temperature in the uniform sector $0 \leqslant r \leqslant a$, $0 \leqslant \varphi \leqslant \alpha$. The temperature on the straight sections of the boundary is zero, and on the arc of the circumference a linear distribution is specified.

Hint: The problem comes down to solving the equation $\Delta u = 0$ in the sector with the boundary conditions $u|_{\varphi=0} = u|_{\varphi=\alpha} = 0$, $u_{r=a} = A\varphi$ (A = const).

Answers

1. $u(r, \varphi) = 2 + 3\frac{r}{r_0}\sin\varphi$. 2. $u(r, \varphi) = \frac{1}{2} - \frac{1}{2}\left(\frac{r}{r_0}\right)^2 \cos 2\varphi$. 3. $u = A_0 + ar\cos\varphi$, A_0 is an arbitrary constant. 4. The problem has no solution. 5. $u(r, \varphi) = 1 - \frac{\ln r}{\ln 2} + \left(\frac{r}{3} - \frac{4}{3r}\right)\cos\varphi + \frac{4}{15}\left(r^2 - \frac{1}{r^2}\right)\sin 2\varphi$. 6. $u(r, \varphi) = \frac{\alpha A}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k} \times \left(\frac{r}{a}\right)^{\frac{k\pi}{\alpha}} \sin\frac{k\pi}{\alpha}\varphi$.

Appendix II

Conformal Mappings

In applications it is often necessary to transform a given domain into a simpler domain so that the angles between the curves remained unchanged. Such transformations enable us to solve problems in aero- and fluid-dynamics, elasticity theory, the theory of fields of all sorts, and so on.

We will restrict ourselves to transformations of plane domains.

A continuous mapping $w = f(z)$ of a plane domain D into a domain in a plane is called *conformal at a point* $z_0 \in D$, if at the point it has the property that extensions and angles are the same.

Open domains D_1 and D_2 are said to be *conformally equivalent*, if there exists a one-to-one mapping of one domain onto the other such that it is conformal at each point.

Riemann mapping theorem. *Any two plane open simply connected domains, whose boundaries consist of more than one point, are conformally equivalent.*

In dealing with specific problems the main task is to construct from given plane domains an *explicit* one-to-one conformal mapping of one of them onto another. In the plane case one way is to lean on the tools of the theory of functions of a complex variable.

It has already been noted above (Sec. 26.1) that a univalent analytic function with a nonzero derivative conformally maps its domain on its image.

In constructing conformal mappings the following rule is quite useful:

Principle of boundary correspondence. Suppose that in a simply connected domain D of the complex plane z bounded by a contour γ a function $w = f(z)$ is specified that is continuous on the closure $\overline{D}$ and maps the contour γ onto some contour γ' in the complex plane w. If the sense of tracing the contour is the same, the functon $w = f(z)$ is a conformal mapping of the domain D in the complex plane z onto a domain D' of the complex plane w bounded by the contour γ' (Fig. II.1).

We here wish, using the domains where basic elementary functions of a complex variable are univalent (see Sec. 26.2), to learn to construct con-

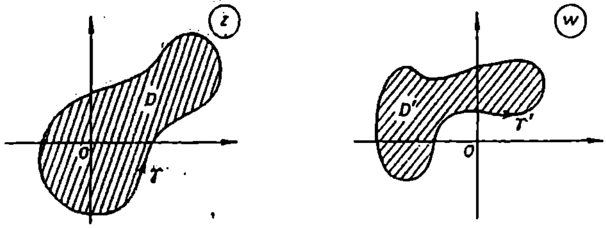


Fig. II.1

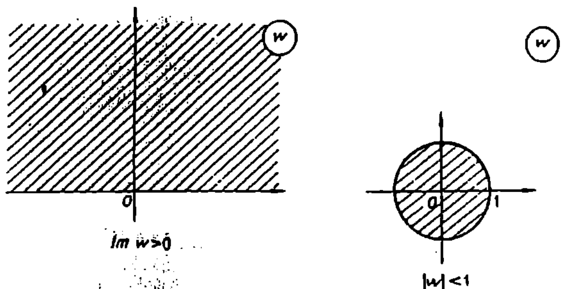


Fig. II.2

formal mappings of open simply connected plane domains, which commonly occur in applications, onto two standard regions—*the upper half-plane* and *unit circle* (Fig. II.2).

In order that the table given below be better used, we provide some simple transformations of the complex plane.

Plane transformations: (1) Parallel translation (shift by a predetermined compex number a (Fig. II.3)).

(2) Rotation (by a predetermined angle φ) (Fig. II.4).

(3) Extension ($k > 1$) and contraction ($0 < k < 1$) (Fig. II.5).

Therefore, a transformation of the type $w = az + b$, $a \neq 0$, can make any circle into a unit circle with centre at zero, and any half-plane can be

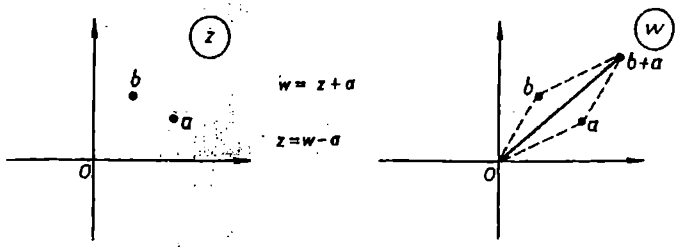


Fig. II.3

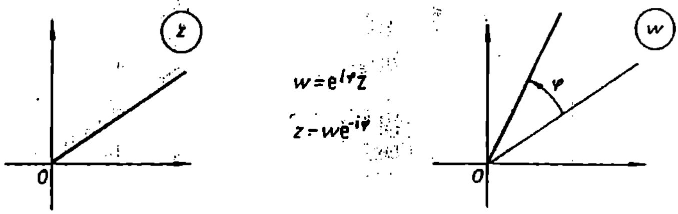


Fig. II.4

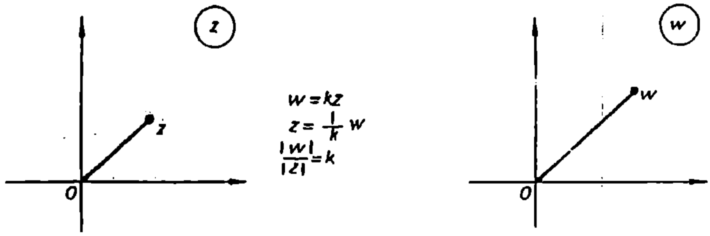


Fig. II.5

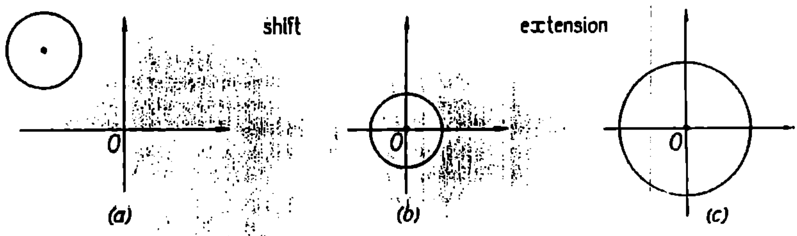


Fig. II.6

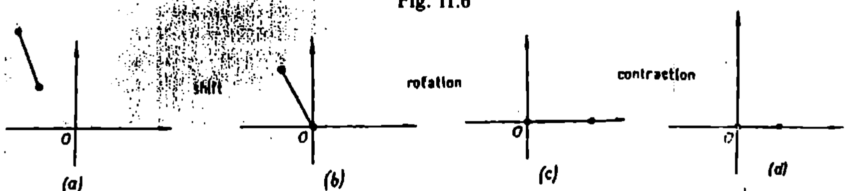


Fig. II.7

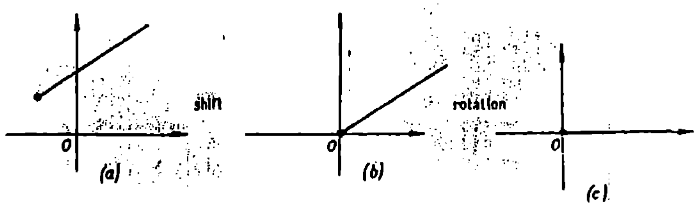


Fig. II.8

made into the upper half-plane, any segment of a straight line can be made into the segment [0, 1] on the real axis, and any ray into the positive x-axis (see Fig. II.6-8).

(4) A transformation of the plane z such that three different points z_1, z_2, z_3 go into different point w_1, w_2, w_3 in the plane (Fig. II.9)

$$\frac{w - w_1}{w - w_2}\,\frac{w_3 - w_2}{w_3 - w_1} = \frac{z - z_1}{z - z_2}\,\frac{z_3 - z_2}{z_3 - z_1}.$$

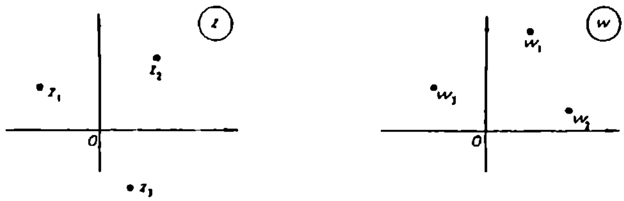


Fig. II.9

Table structure and directions for use. The table is based on the following scheme: serial number, domain D in the complex plane z, conformal mapping (direct $w = f(z)$ and inverse $z = g(w)$), domain D' in the complex plane w, conformally equivalent to D.

Each entry in the table is as a rule either the upper half-plane or the unit circle with centre at zero. As we will see later, such a standardization is convenient in practical applications. For the most part, a transformation is only provided that reduces a given domain to the one considered earlier. In that case, a reference is made to the transformation that transforms the resultant domain into a standard one (unit circle with centre at zero or the upper half-plane).

Table II.1 Basic Elementary Functions

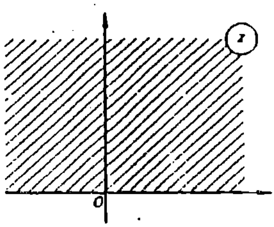
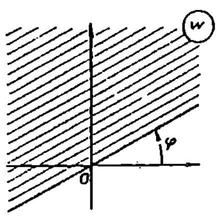
$\operatorname{Im} z > 0$	No. 1	$\varphi < \arg w < \pi + \varphi$
	$w = e^{i\varphi} z$ $z = e^{-i\varphi} w$	

Table 11.1 (continued)

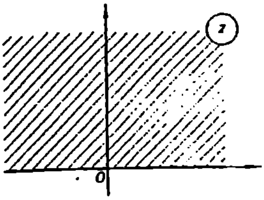
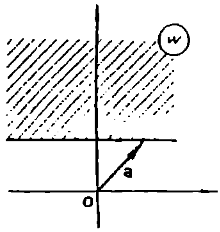
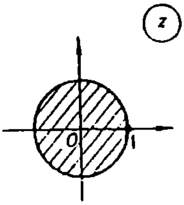
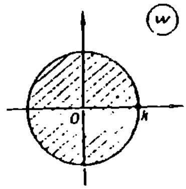
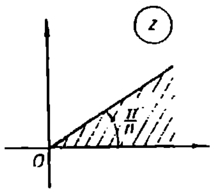
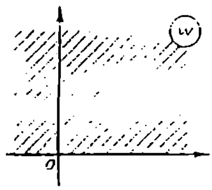
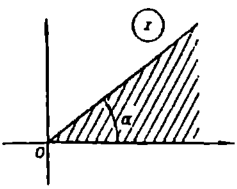
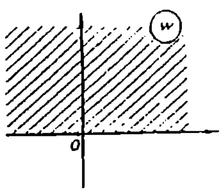
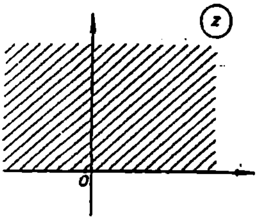
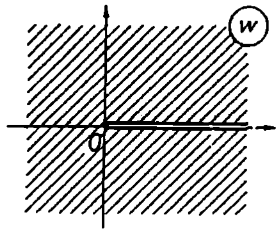
$\operatorname{Im} z > 0$	No. 2	$\operatorname{Im} w > \operatorname{Im} a$
	$w = z + a$ $z = w - a$	
$\lvert z\rvert < 1$	No. 3	$\lvert w\rvert < k$
	$w = kz$ $z = \frac{1}{k} w$ $k > 1$	
$0 < \arg z < \frac{\pi}{n}$	No. 4	$\operatorname{Im} w > 0$
	$w = z^n$ $z = \sqrt[n]{w}$ $n = 2, 3, \ldots$	
$0 < \arg z < \alpha < \pi$	No. 5	$\operatorname{Im} w > 0$
	$w = z^{\frac{\pi}{\alpha}}$ $z = w^{\frac{\alpha}{\pi}}$	

Table II.1 (continued)

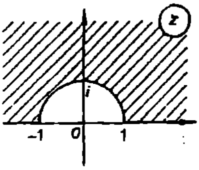
$\operatorname{Im} z > 0$	No. 6	Ray $[0, +\infty)$ on the real axis is removed from the plane w



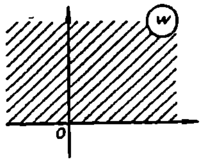
$$w = z^2$$
$$z = \sqrt{w}$$



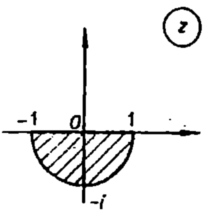
$\operatorname{Im} z > 0, \quad \lvert z\rvert > 1$	No. 7	$\operatorname{Im} w > 0$



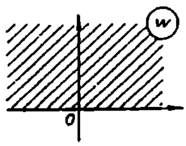
$$w = \frac{1}{2}\left(z + \frac{1}{z}\right)$$



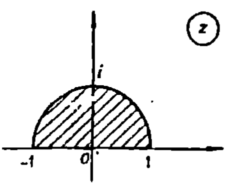
$\operatorname{Im} z < 0, \quad \lvert z\rvert < 1$	No. 8	$\operatorname{Im} w > 0$



$$w = \frac{1}{2}\left(z + \frac{1}{z}\right)$$



$\operatorname{Im} z > 0, \quad \lvert z\rvert < 1$	No. 9	$\operatorname{Im} w > 0$



$$w = -\frac{1}{2}\left(z + \frac{1}{z}\right)$$

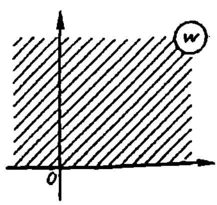
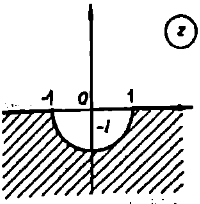
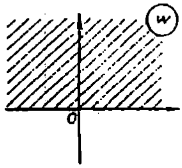


Table II.1 (concluded)

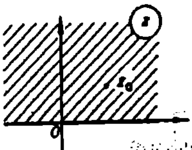
$\operatorname{Im} z < 0,\ |z| > 1$ — No. 10 — $\operatorname{Im} w > 0$



$$w = -\frac{1}{2}\left(z + \frac{1}{z}\right)$$

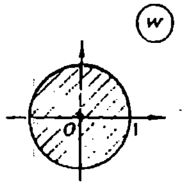


$\operatorname{Im} z > 0$ — No. 11 — $|w| < 1$

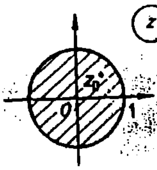


$$w = e^{i\alpha}\frac{z - z_0}{z - \bar{z}_0}$$

$$w(z_0) = 0$$



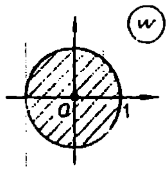
$|z| < 1$ — No. 12 — $|w| < 1$



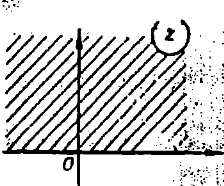
$$w = e^{i\alpha}\frac{z - z_0}{1 - \bar{z}_0 z}$$

$$w(z_0) = 0$$

$$\arg w'(z_0) = \alpha$$



$\operatorname{Im} z > 0$ — No. 13 — $\operatorname{Im} w > 0$



$$w = \frac{az + b}{cz + d}$$

a, b, c, d are real numbers

$ad - bc > 0.$

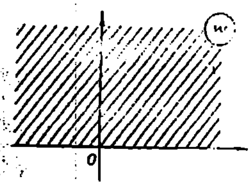


Table II.2 Plane with Cuts

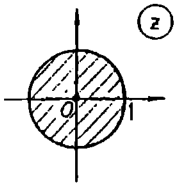
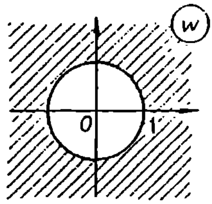
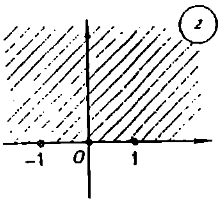
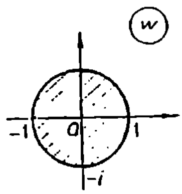
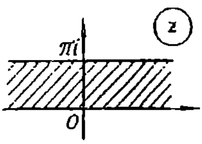
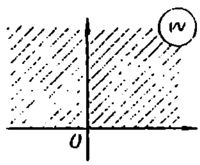
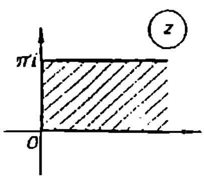
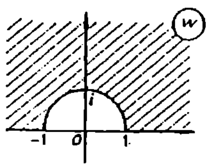
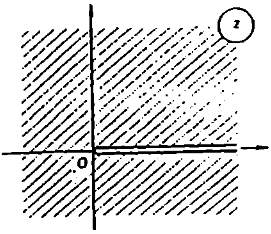
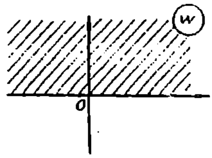
$\lvert z\rvert < 1$	No. 14	$\lvert w\rvert > 1$
	$w = \frac{1}{z}$ $z = \frac{1}{w}$	
$\operatorname{Im} z > 0$	No. 15	$\lvert w\rvert < 1$
	$\frac{w+1}{w+i}\cdot\frac{1+i}{1+1}$ $= \frac{z+1}{z-0}\,\frac{1-0}{1+1}$	
$0 < \operatorname{Im} z < \pi$	No. 16	$\operatorname{Im} w > 0$
	$w = e^z$ $z = \ln w$	
$0 < \operatorname{Im} z < \pi, \quad \operatorname{Re} z > 0$	No. 17	$\operatorname{Im} w > 0, \quad \lvert w\rvert > 1$
	$w = e^z$ $z = \ln w$	

Table 11.2 (continued)

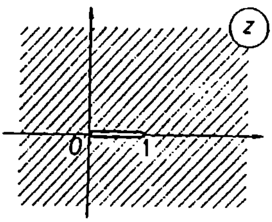
Plane cut along the ray $[0, +\infty)$	No. 18	$\operatorname{Im} w > 0$



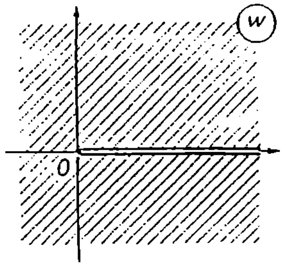
$$w = \sqrt{z}$$
$$z = w^2$$



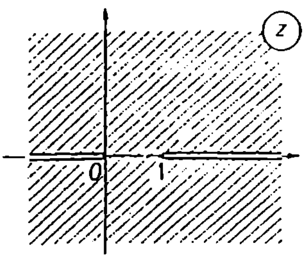
Plane cut along the segment $[0, 1]$	No. 19	Plane cut along the ray $[0, +\infty)$



$$w = \frac{z}{1 - z}$$
$$z = \frac{w}{w + 1}$$



Plane with two cuts $(-\infty, 0]$ and $[1, +\infty)$	No. 20	Plane cut along the ray $[0, +\infty)$



$$w = \frac{z}{z - 1}$$
$$z = \frac{w}{w - 1}$$

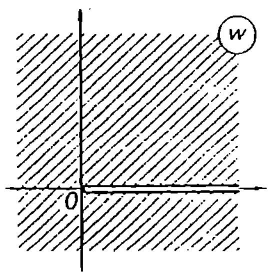
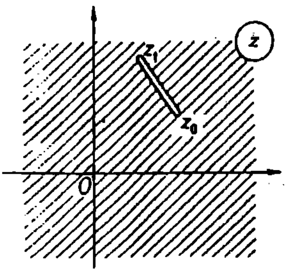


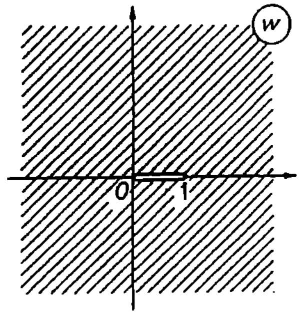
Table II.2 (continued)

Plane cut along the segment $[z_0, z_1]$	No. 21	Plane cut along the segment $[0, 1]$

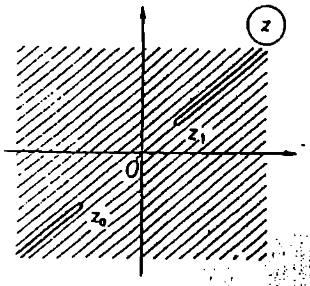


$$w = \frac{z - z_0}{z_1 - z_0}$$

$$z = z_0 + (z_1 - z_0)w$$

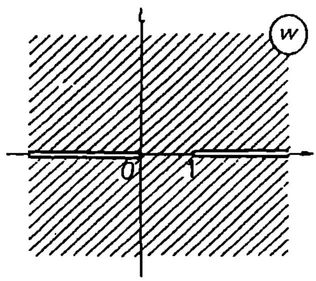


Plane cut along two rays lying on one straight line	No. 22	Plane cut along the rays $(-\infty, 0]$ and $[1, +\infty)$

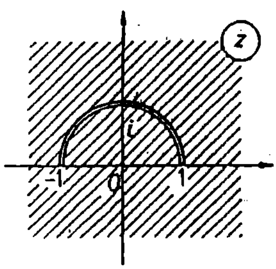


$$w = \frac{z - z_0}{z_1 - z_0}$$

$$z = z_0 + (z_1 - z_0)w$$



Plane cut along the semicircle $\lvert z\rvert = 1$, $\operatorname{Im} z > 0$	No. 23	Plane cut along the ray $[0, +\infty)$



$$w = \frac{1}{i}\,\frac{z - 1}{z + 1}$$

$$z = \frac{iw + 1}{1 - iw}$$

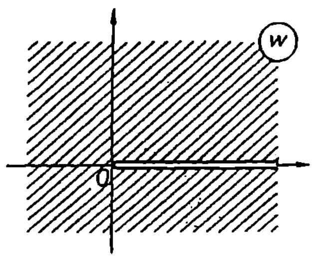
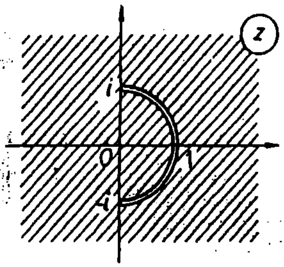


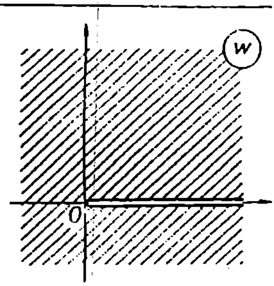
Table II.2 (concluded)

Plane cut along the semicircle $\lvert z\rvert = 1$, $\mathrm{Re}\, z > 0$	No. 24	Plane cut along the ray $[0, +\infty)$

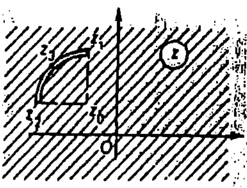


$$w = i\frac{z - i}{z + i}$$

$$z = \frac{w + i}{1 + iw}$$



Plane cut along the arc of a circle $\lvert z - z_0\rvert = r$, $z = z_0 + re^{i\varphi}$, $\varphi_1 \leqslant \varphi \leqslant \varphi_2$	No. 25	Plane cut along the ray $[0, +\infty)$



$z_1 = z(\varphi_1)$, $z_2 = z(\varphi_2)$

$$z_3 = z\left(\frac{\varphi_1 + \varphi_2}{2}\right)$$

$$w = \frac{z_3 - z_2}{z_3 - z_1}\,\frac{z - z_1}{z - z_2}$$

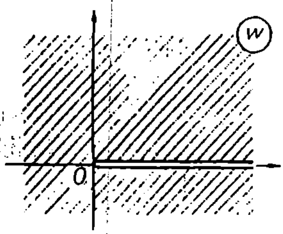
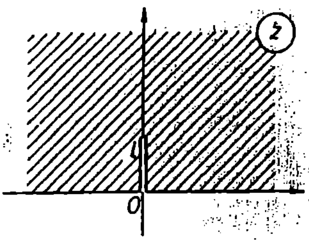


Table II.3 Half-Plane with Cuts

Half-plane $\mathrm{Im}\, z > 0$ with a cut along the segment $[0, i]$	No. 26	Plane with a cut along the ray $[-1, +\infty)$



$$w = z^2$$

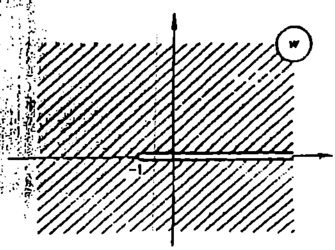


Table II.3 (concluded)

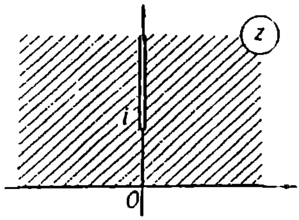
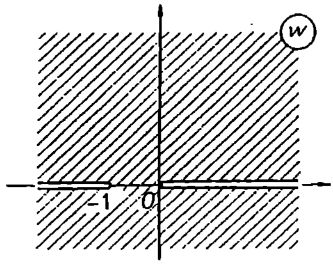
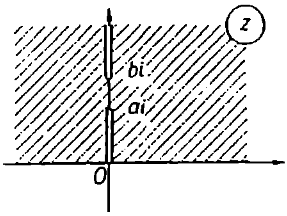
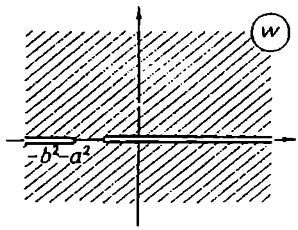
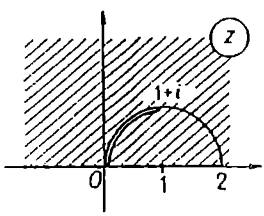
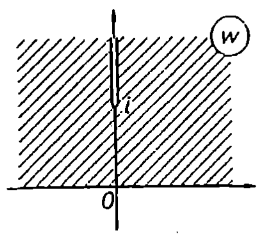
Half-plane $\operatorname{Im} z > 0$ with a cut along the ray $[i, i\infty)$ of the imaginary axis	No. 27	Plane with cuts along the rays $(-\infty, -1]$ and $[0, +\infty)$
	$w = z^2$	
Half-plane $\operatorname{Im} z > 0$ with cuts along the ray $(bi, +i\infty)$ and the segment $[0, ai]$ of the imaginary axis, $0 < a < b$	No. 28	Plane with cuts along the rays $(-\infty, -b^2]$ and $[-a^2, +\infty)$
	$w = z^2$	
Half-plane with a cut along the arc of the circle $\|z - 1\| = 1$, $z = 1 + e^{i\varphi}$, $\frac{\pi}{2} \leqslant \varphi \leqslant \pi$	No. 29	Half-plane $\operatorname{Im} w > 0$ with a cut along the ray $[i, i\infty)$ of the imaginary axis
	$w = \dfrac{z - 1}{z}$ $z = \dfrac{1}{1 - w}$	

Table II.4 Circle with Cuts

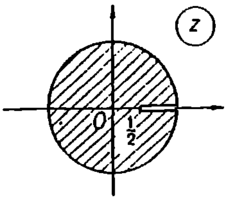
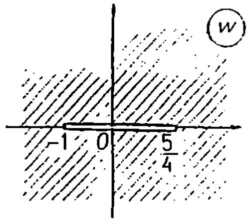
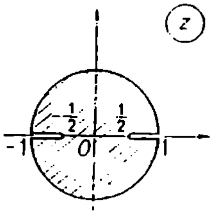
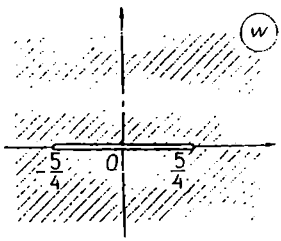
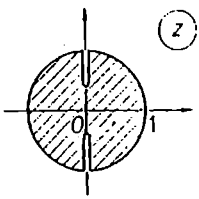
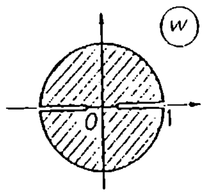
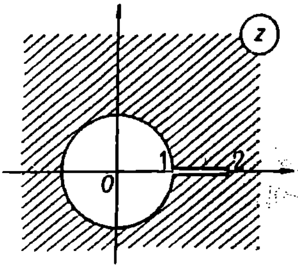
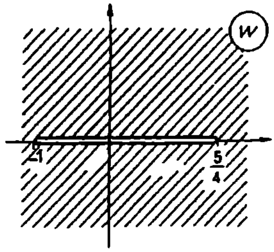
Circle $\|z\| < 1$ with a cut along the radius: $[1/2, 1]$	No. 30	Plane with a cut along the segment $[-1, 5/4]$
	$w = \frac{1}{2}\left(z + \frac{1}{z}\right)$	
Circle $\|z\| < 1$ with two cuts along the diameter on the real axis	No. 31	Plane with a cut along the segment $[-5/4, 5/4]$
	$w = \frac{1}{2}\left(z + \frac{1}{z}\right)$	
Circle $\|z\| < 1$ with two cuts along the diameter on the imaginary axis	No. 32	Circle $\|w\| < 1$ with two cuts along the diameter on the real axis
	$w = \frac{z}{i}$	

Table II.5 The Outside of a Circle with Cuts

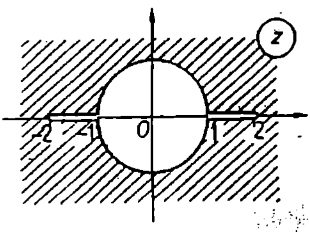
The outside of the circle $|z| > 1$ with a cut along the segment $[1, 2]$	No. 33	Plane with a cut along the segment $[-1, 5/4]$



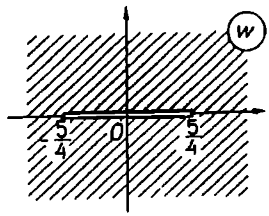
$$w = \frac{1}{2}\left(z + \frac{1}{z}\right)$$



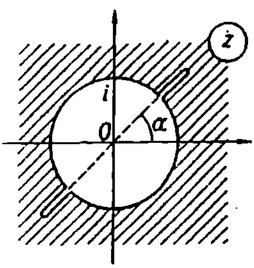
The outside of the circle $|z| > 1$ with two cuts along the segments $[-2, -1]$ and $[1, 2]$	No. 34	Plane with cuts along the segment $[-5/4, 5/4]$



$$w = \frac{1}{2}\left(z + \frac{1}{z}\right)$$



The outside of the circle $|z| > 1$ with cuts along two segments of a straight line	No. 35	The outside of the circle $|w| > 1$ with cuts along two segments of the real axis



$$w = e^{-i\alpha} z$$

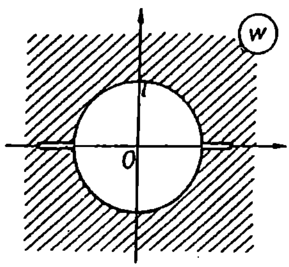
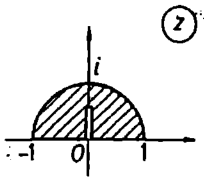


Table II.6 Semicircle with Cuts

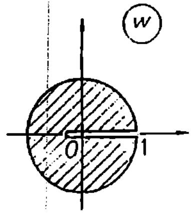
Semicircle $|z| < 1$, $\operatorname{Im} z > 0$ with a cut along the segment $[0, i/2]$ on the imaginary axis

No. 36

Circle $|w| < 1$ with a cut along the segment $[-1/4, 1]$ on the real axis



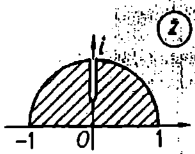
$w = z^2$



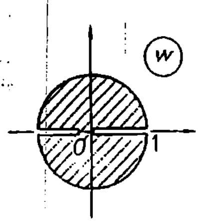
Semicircle $|z| < 1$, $\operatorname{Im} z > 0$ with a cut along the segment $[i/2, i]$ on the imaginary axis

No. 37

Circle $|w| < 1$ with cuts along the segments $[-1, -1/4]$ and $[0, 1]$ on the real axis



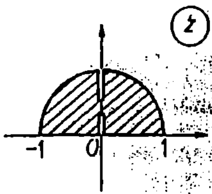
$w = z^2$



Semicircle $|z| < 1$, $\operatorname{Im} z > 0$ with two cuts along the segments $[0, ai]$ and $[bi, i]$ on the imaginary axis; $0 < a < b < 1$

No. 38

Circle $|w| < 1$ with cuts along the segments $[-1, -b^2]$ and $[-a^2, 1]$ on the real axis



$w = z^2$

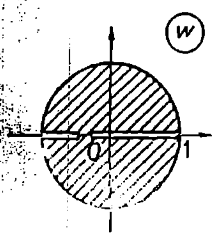
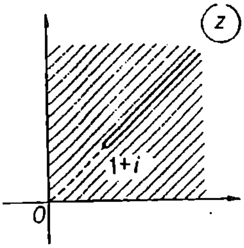
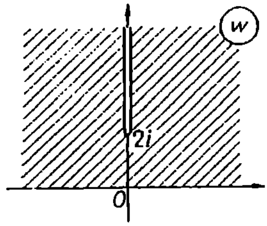


Table II.7 Angle with Cuts

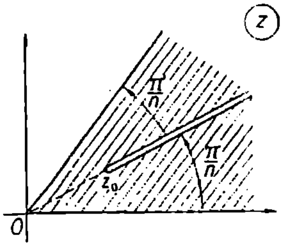
Angle $0 < \arg z < \pi/2$ with a cut along the ray $\arg z = \pi/4$ originating at $1 + i$	No. 39	Half-plane $\operatorname{Im} w > 0$ with a cut along the ray on the imaginary axis originating at $2i$



$w = z^2$



Angle $0 < \arg z < 2\pi/n$ with a cut along the ray $\arg z = \pi/n$ with origin at $z_0 = e^{i\frac{\pi}{n}}a$; $a > 0$	No. 40	Plane with two cuts along the rays $(-\infty, -a^n]$ and $[0, +\infty)$



$w = z^n$

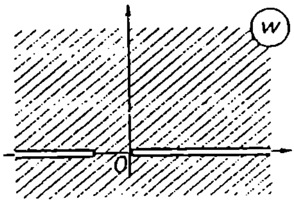
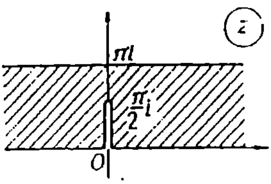


Table II.8 Band with Cuts

Band $0 < \operatorname{Im} z < \pi$ with a cut along the segment $[0, i\pi/2]$ on the imaginary axis	No. 41	Half-plane $\operatorname{Im} w > 0$ with a cut along the arc $w = e^{i\varphi}$, $0 \leqslant \varphi \leqslant \pi/2$



$w = e^z$

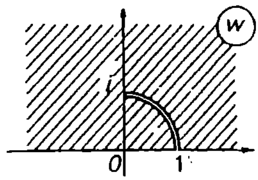
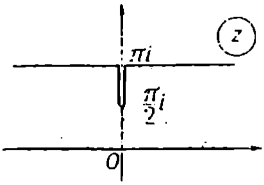
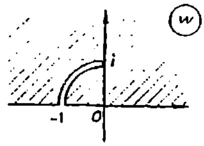


Table 11.8 (continued)

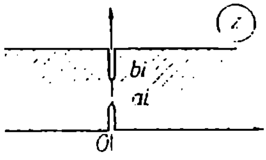
Band $0 < \operatorname{Im} z < \pi$ with a cut along the segment $[\pi i/2, \pi i]$ on the imaginary axis	No. 42	Half-plane $\operatorname{Im} w > 0$ with a cut along the arc $w = e^{i\varphi}$, $\pi/2 \leqslant \varphi \leqslant \pi$



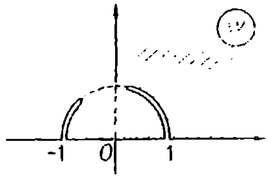
$w = e^z$



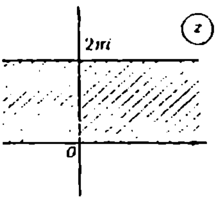
Band $0 < \operatorname{Im} z < \pi$ with two cuts along the segments $[0, ai]$ and $[bi, \pi i]$, $0 < a < b < \pi$, on the imaginary axis	No. 43	Half-plane with cuts along the arcs $w = e^{i\varphi}$, $0 \leqslant \varphi \leqslant a$, $b \leqslant \varphi \leqslant \pi$



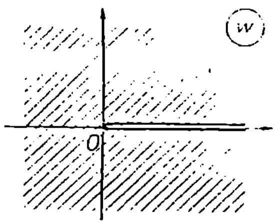
$w = e^z$



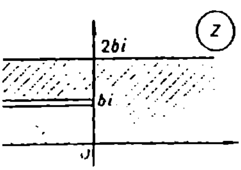
Band $0 < \operatorname{Im} z < 2\pi$	No. 44	Plane with a cut along the ray $[0, +\infty)$ on the real axis



$w = e^z$



Band $0 < \operatorname{Im} z < 2b$, $b > 0$ with a cut along the ray $\operatorname{Im} z = b$, $\operatorname{Re} z \leqslant 0$	No. 45	Half-plane $\operatorname{Im} z > 0$ with a cut along the segment $[0, ai]$ on the imaginary axis



$w = e^{\frac{\pi z}{2b}}$

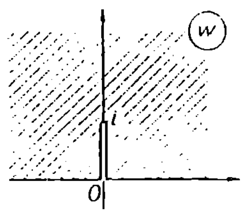
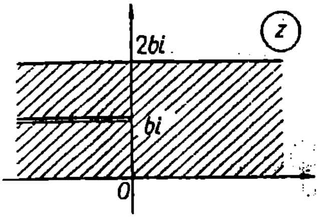
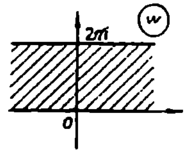


Table II.8 (concluded)

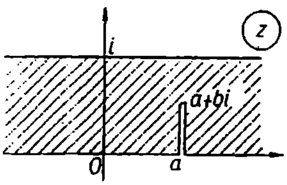
Band $0 < \operatorname{Im} z < 2b$, $b > 0$ with a cut along the ray $\operatorname{Im} z = b$, $\operatorname{Re} z \leqslant 0$	No. 46	Band $0 < \operatorname{Im} z < 2\pi$



$$w = \ln\left(e^{\frac{\pi z}{b}} + 1\right)$$



Band $0 < \operatorname{Im} z < 1$ with a cut along the segment $\operatorname{Re} z = a$, $0 \leqslant \operatorname{Im} z \leqslant b < 1$	No. 47	Band $0 < \operatorname{Im} w < 1$



$$w = \sqrt{\tanh^2 \frac{\pi(z - a)}{2} + \tan^2 \frac{\pi b}{2}}$$

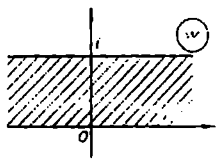
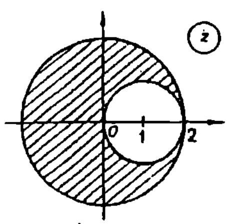


Table II.9 Mappings of Excized Regions

Circular crescent $\lvert z\rvert < 2$, $\lvert z - 1\rvert > 1$	No. 48	Band $0 < \operatorname{Im} w < 1$



$$w = 2i\,\frac{z}{z - 2}$$

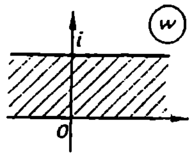


Table II.9 (continued)

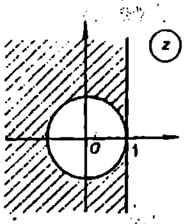
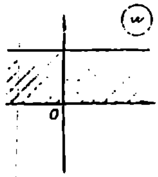
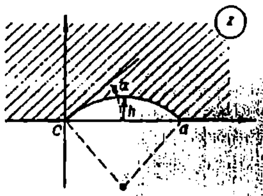
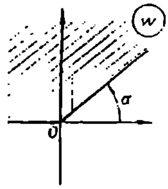
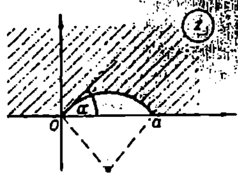
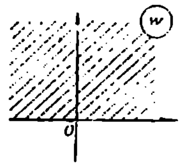
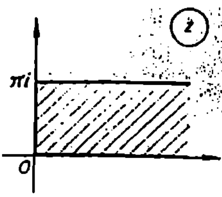
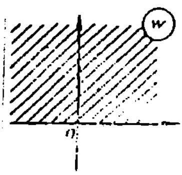
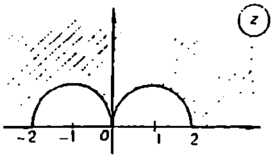
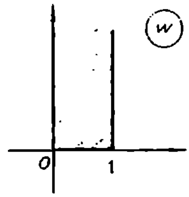
Region $\operatorname{Re} z < 1$, $\lvert z\rvert > 1$	No. 49	Band $0 < \operatorname{Im} w < 1$
	$w = i\dfrac{z+1}{z-1}$	
Half-plane $\operatorname{Im} z > 0$ with cut-out segments of a circle	**No. 50**	**Angle $\alpha < \arg w < \pi$**
	$w = \dfrac{z}{a-z}$	
Half-plane $\operatorname{Im} z > 0$ with a cut-out segment of a circle	**No. 51**	**Half-plane $\operatorname{Im} w > 0$**
	$w = -\left(\dfrac{z}{z-a}\right)^{\frac{\pi}{\pi-\alpha}}$	
Half-band $0 < \operatorname{Im} z < \pi$, $\operatorname{Re} z > 0$	**No. 52**	**Half-plane $\operatorname{Im} w > 0$**
	$w = \cosh z$	

Table II.9 (continued)

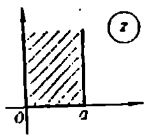
Half-plane $\operatorname{Im} z > 0$ with cut-out semicircles — No. 53 — Half-band $0 < \operatorname{Re} z < 1$, $\operatorname{Im} z > 0$



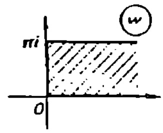
$$w = \frac{z - 2}{z}$$



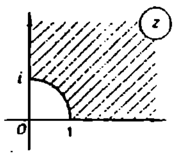
Half-band $0 < \operatorname{Re} z < a$, $\operatorname{Im} z > 0$ — No. 54 — Half-band $0 < \operatorname{Im} w < \pi$, $\operatorname{Re} w > 0$



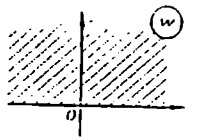
$$w = \pi \frac{\dfrac{z}{a} - 1}{i}$$



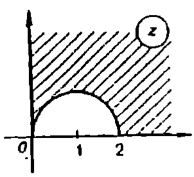
Angle $\operatorname{Im} z > 0$, $\operatorname{Re} z > 0$ with a cut-out sector $|z| < 1$ — No. 55 — Half-plane $\operatorname{Im} w > 0$



$$w = \frac{1}{2}\left(z^2 + \frac{1}{z^2}\right)$$



Angle $\operatorname{Im} z$, $\operatorname{Re} z > 0$ with a cut-out semicircle — No. 56 — Half-band $0 < \operatorname{Re} w < 1$, $\operatorname{Im} w > 0$



$$w = -\frac{2}{z}$$

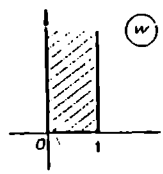


Table 11.9 (continued)

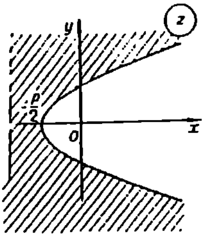
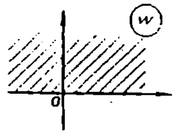
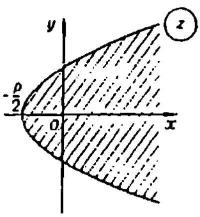
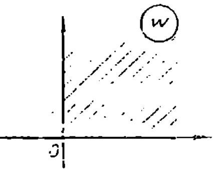
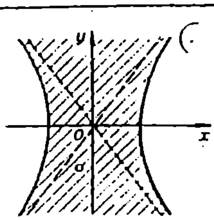
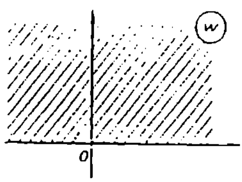
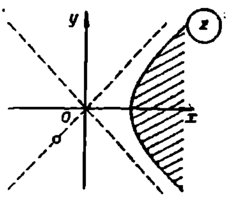
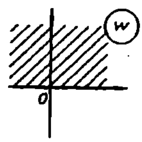
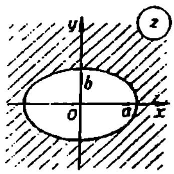
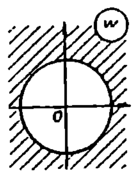
The outside of the parabola $y^2 = 2p\left(x + \frac{p}{2}\right)$, $p > 0$, $z = x + iy$	No. 57	Half-plane $\operatorname{Im} w > 0$
	$w = \sqrt{z} - i\sqrt{\frac{p}{2}}$	
Inside of the parabola $y^2 = 2p\left(x + \frac{p}{2}\right)$, $p > 0$ $z = x + iy$	No. 58	Half-plane $\operatorname{Im} w > 0$
	$w = i\sqrt{2}\cosh \pi\sqrt{\frac{z}{2p}}$	
Outside of the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, $c = \sqrt{a^2 + b^2}$, $\theta = \sin^{-1}\frac{a}{c}$	No. 59	Half-plane $\operatorname{Im} w > 0$
	$w = \left(\frac{z + \sqrt{z^2 - c^2}}{ce^{i\theta}}\right)^{\frac{\pi}{\pi - 2\theta}}$	

Table II.9 (concluded)

Inside of the right-hand branch of the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, $c = \sqrt{a^2 + b^2}$, $\theta = \sin^{-1}\frac{a}{c}$	No. 60	Half-plane $\operatorname{Im} w > 0$
	$w = i\sqrt{2}\cosh\left(\frac{\pi}{2\theta}\cosh^{-1}\frac{z}{c}\right)$	
Outside of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $a > b > 0$, $c = \sqrt{a^2 - b^2}$	No. 61	Outside of the circle $\lvert w\rvert > 1$
	$w = \frac{z + \sqrt{z^2 - c^2}}{a + b}$	

We will now illustrate by an example how the table is to be used.

Example. Find a one-to-one and conformal mapping of the circle

$$|z - (2 + 2i)| < 3$$

with a cut along the radius (Fig. II.10)

$$\arg z = \frac{\pi}{4}, \qquad 1 + 2\sqrt{2} \leqslant |z| < 3 + 2\sqrt{2}$$

onto the unit circle with centre at zero.

◄ (a) We will make use of the elementary plane trasformations to reduce the given domain to the one available in the table.

1. We shift the centre of the given circle to zero (see Fig. II.11)

$$z_1 = z - (2 + 2i).$$

We have: circle $|z_1| < 3$ with a cut $\arg z_1 = \pi/4$, $1 \leqslant |z_1| < 3$.

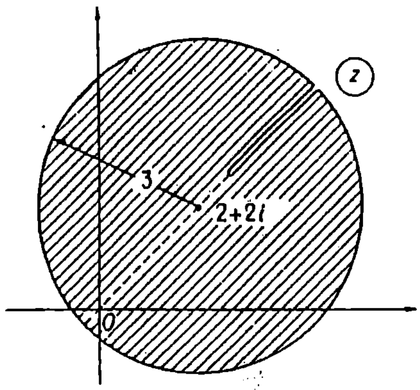


Fig. II.10

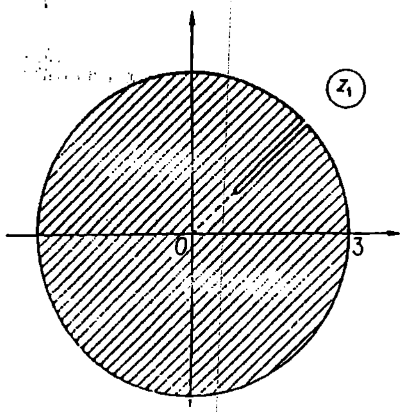


Fig. II.11

2. We turn the resultant circle clockwise by the angle (Fig. II.12)

$$z_2 = z_1 e^{-i\frac{\pi}{4}} = z_1 \frac{\sqrt{2} - i\sqrt{2}}{2}.$$

We have: circle $|z_2| < 3$ with a cut $\arg z_2 = 0$, $1 \leqslant |z_2| < 3$.

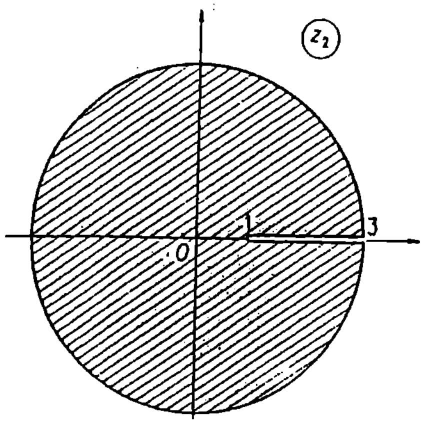


Fig. II.12

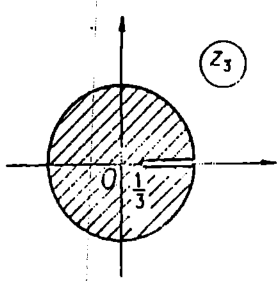


Fig. II.13

3. We compress the circle so that (Fig. II.13)

$$z_3 = \frac{z_2}{3}.$$

We have: circle $|z_3| < 1$ with a cut $1/3 \leqslant |z_3| < 1$, $\arg z_3 = 0$.

The given domain is thus reduced to a tabulated one using the transformation

$$z_3 = \frac{z - 2 - 2i}{3\sqrt{2}}(1 - i).$$

(b) 1. The given region is a circle $|z_3| < 1$ with a cut $\arg z_3 = 0$, $1/3 \leqslant |z_3| < 1$. It is given in the table (No. 30). The Joukowsky function

$$z_4 = \frac{1}{2}\left(z_3 + \frac{1}{z_3}\right)$$

transforms this region into a plane with a cut along the segment $[-1, 5/3]$ on the real axis (Fig. II.14).

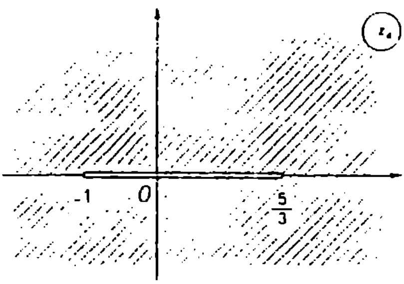


Fig. II.14

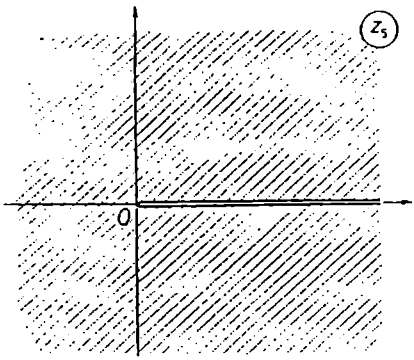


Fig. II.15

2. This domain is given in the table (No. 22). Using the linear fractional transformation

$$z_5 = \frac{z_4 + 1}{\frac{5}{3} - z_4},$$

we transform the region into a plane with a cut along the ray $[0, +\infty)$ on the real axis (Fig. II.15).

3. This region is given in the table (No. 6). Extracting the root

$$z_6 = \sqrt{z_5},$$

we transform this region into the upper half-plane $\operatorname{Im} z_6 > 0$ (Fig. II.16).

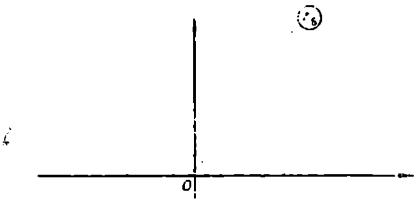


Fig. II.16

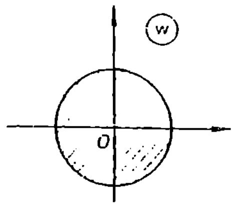


Fig. II.17

4. This region is given in the table (No. 11). Using the linear fractional transformation

$$w = z_7 = \frac{z_6 - 1}{1 - iz_6}$$

we transform the region into the unit circle with centre at zero (Fig. II.17).

$$|w| < 1.$$

By successively expressing z_k through z_{k-1}, we will obtain a one-to-one conformal mapping of the given circle with a cut along the radius defined on the complex plane z onto the unit circle on the complex plane w. ►

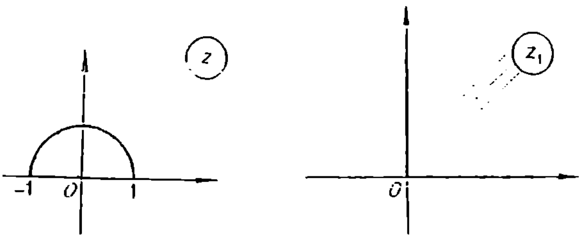


Fig. II.18

Conformal mapping is not uniquelly defined by given domains.

Example. Find a one-to-one and conformal mapping of the semicircle

$$|z| < 1, \qquad \operatorname{Im} z > 0$$

on the upper half-plane

$$\operatorname{Im} w > 0.$$

◄ (a) 1. The linear fractional mapping

$$z_1 = \frac{z + 1}{1 - z}$$

transforms the given semi-circle into a right angle (Fig. II.18)

$$\operatorname{Im} z_1 > 0, \qquad \operatorname{Re} z_1 > 0.$$

2. The domain is given in the table (No. 4, $n = 2$). Raising to the square

$$w = z_1^2,$$

we map this domain onto the upper half-plane $\operatorname{Im} w > 0$. Thereby, the desired mapping is

$$w = \left(\frac{z+1}{z-1}\right)^2.$$

(b) The given domain

$$|z| < 1, \qquad \operatorname{Im} z > 0$$

is given in the table (No. 9).

The required mapping has the form

$$w = -\frac{1}{2}\left(z + \frac{1}{z}\right).$$

Both mappings

$$w = \left(\frac{z+1}{z-1}\right)^2 \quad \text{and} \quad w = -\frac{1}{2}\left(z + \frac{1}{z}\right)$$

are one-to-one conformal mappings of the given semicircle $|z| < 1$, $\operatorname{Im} z > 0$ onto the upper half-plane

$$\operatorname{Im} w > 0. \quad ▶$$

Index